

環境取樣學課程內容綱要	<605ENV-SAMPLINGintroduction.doc>
高等森林測計學課程內容綱要	<adv-msu-intro.doc>
第一章 前言(高義盛)	<sam-ch01.doc>
第八章 迴歸推估(高義盛)	<sam-ch08.doc>
第十二章 群團和系統取樣(高義盛)	<sam-ch12.doc>
第十三章 多層級取樣設計(高義盛)	<sam-ch13.doc>
第十四章 雙重取樣(高義盛)	<sam-ch14.doc>
第二十六章 分層適應群團取樣(高義盛)	<sam-ch26.doc>
地理資訊系統於分層取樣調查設計與分析上之應用	<SAM-DEZ.doc>
植群分析樣區大小與形狀之探討(羅南璋)	<samplesize&shape-ch12.doc>
資源調查樣本數決定之探討	<SAMP-NO.doc>
森林調查法	<Samp-tech.doc>
地理資訊系統於分層取樣調查設計與分析上之應用	<Sam-stratfied.doc>
地理資訊系統於分層取樣調查設計與分析上之應用	<SAM-stratfied.doc>
森林資源取樣調查方法之研究	<TW-FIA-SAMP.doc>
環境取樣學目錄研讀(高義盛)	<環境取樣學.doc>

【教師姓名】 馮 豐 隆 【選修系級】 研究所一年級以上學生可選修
【科目代號】 【科目中文名稱】 環境取樣學
 【科目英文名稱】 Environmental Sampling
【必修或選修】 選修 【學分數】 3 學分，上學期（ — ） 下學期（ ）
【講授】 3 小時 【實習】 小時 【先修科目】

【教學目的】

討論在生態系經營理念下，資料如何收集、整理、分析與模式建立，以為提供森林多目標經營管理決策所需之資訊。課程中，尤其集中於生態系經營目標之確定，資料收集方式，取樣調查方法，整合調查資料處理、分析與模式建立。有關統計模式、模擬模式、專家系統應用在(a)森林資源調查、生長與收穫之理論與(b)森林資源經營之決策支援系統(DSS) 建造技術、理論與論文的研讀。

【內容綱目】

1. 論文的查詢、收集、研讀、分析、整理與批判。
2. 生態系經營、地景經營與森林多目標經營管理與其經營管理時考慮之資訊及其調查項目。
3. 森林資源調查學(取樣調查、地面調查、航測、遙測學)的趨勢，逢機、分層、雙重、雙層、多層、樣點、樣線、截線、空間、適應取樣調查方法與實務。
4. 樣區 — 永久樣區(調查、生長收穫樣區)、臨時樣區收集資料、CFI與整合森林資源調查資料，多尺度調查資料之整合處理、分析與建造，及方法、技術、理論發展的趨勢。
5. 討論資源科技(遙測、電腦、網路、自動化、GIS、GPS、ES、WWW....)在(a)森林資源調查(b)生長與演替之理論與(c)森林資源經營監測上之應用。
6. 討論統計模式、模擬模式、專家系統在(a)森林資源調查與監測(b)生長與演替上之應用。
7. 收集、分析具時間(Temporal—status, dynamics)、空間(Spatial)之資料有關問題。

授課教師（簽章）_____

一、課程介紹

高等森林測計學(Advanced Forest Measurement)

代 碼： 6055 (3學分)

授課老師：馮豐隆 博士 (森林系)

選課條件：研究所一年級以上學生可選修

課程簡介：討論在生態系經營理念下，資料如何收集、整理、分析與模式建立，以為提供森林多目標經營管理決策所需之資訊。課程中，尤其集中於生態系經營目標之確定，資料收集方式，取樣調查方法，整合調查資料處理、分析與模式建立。有關統計模式、模擬模式、專家系統應用在 (a)森林資源調查、生長與收穫之理論與(b)森林資源經營之決策支援系統(DSS) 建造技術、理論與論文的研讀。模擬

教 材：相關內容的著名參考文獻，見於以下重要雜誌：

1. 日林誌
2. AI Applications in Natural Resource Management
3. Biomathematics
4. BioScience
5. Ecological modelling
6. Forest Science
7. Geo-Info, GIS world, GIS news
8. GPS
9. Journal of Environmental Economics and Management
10. Journal of Management Information System
11. Photogrammetry and Remote Sensing
12. Simulation

課程評分：期中報告佔30%，平時作業佔40%，期末報告佔30%

上課時間：另訂

上課地點：森林館210研究室

二、課程內容

- 內容：1. 論文的查詢、收集、研讀、分析、整理與批判。
2. 生態系經營、景觀經營與森林多目標經營管理與其經營管理時考慮之資訊及其調查項目。
3. 森林資源調查學(取樣調查、地面調查、航測、遙測學)的趨勢，逢機、分層、雙重、雙層、多層、樣點、樣線、截線、空間、適應取樣調查方法與實務。
4. 樣區-- 永久樣區(調查、生長收穫樣區)、臨時樣區收集資料、CFI與整合森林資源調查資料，多尺度調查資料之整合處理、分析與建造，及方法、技術、理論發展的趨勢。
5. 討論資源科技(遙測、電腦、網路、自動化、GIS、GPS、ES...)在 (a)森林資源調查(b)生長與演替之理論與(c)森林資源經營監測上之應用。
6. 討論統計模式、模擬模式、專家系統在 (a)森林資源調查與監測(b)生長與演替上之應用。
7. 收集、分析具時間(Temporal--status,dynamics)、空間(Spatial)之資料有關問題。

整個過程可以以下"自然資源經營管理決策支援系統"架構表示之。

Decision-making in Natural Resource Management

1. Objectives

- (1)People--Economical, Social (Global, National, Local...)
- (2)Environment--Biological, Ecological (Ecosystem, Landscape, Land, Stand...)

2. Data Collection

- (1)Data Determination--survey items, record types, questionnaire
- (2)Edit Survey Guide--survey objectives
 - sampling design (cost, accuracy)
 - methods (permanent plots)
 - What, Where, When, Who ,How(Objectives, Spatial, Temporal, Persons, Technology)
- (3)Data Acquisition (Origin) --
 - A.Exisiting data
 - B.Ground field survey
 - Species identification
 - Items protocol
 - Measurement methodology
 - C.Remote sening
 - Different platforms(scaling)
 - Global, National, Regional, Stand, Tree, Organ (accuracy, calibration)
 - Different sensors
 - photo image, digital image, MSS, video
 - GPS, Distance optical tools be used
 - Photo interpretation
 - Image processing

3. Database Management and Query system

GIS , Database, CD, Visual basic, Mosaic

- (1)Data input
- (2)Data storage
- (3)Data output (Data display)

4.Data Analysis

- (1)Attribute data-statistic description (tables, figures)
- (2)Spatial map layers-- overlapping/neighborhood/link

5.Model Development

- (1)Descriptive models
- (2)Estimation models (projection models)
- (3)Prediction models

6.Decision-making Supporting System

- (1)productivity--FIA, FHMS
- (2)conservation--GAP, NHP (National Heritage Program)

7.Decision-making

Information & decision-making criteria.

Natural Resource management System

- (1)Management Concepts Sustainable
 - Land Classification--Multiple Use
 - Integration--Ecosystem
 - Adaptive Management Area Management
- (2)Management Criteria
 - Ecological --Diversity and Productivity
 - Biological
 - Geographic
 - Social-Economic--Diversity and Productivity
- (3)Management Strategies--Demand, Supply
 - Land Planning
 - Land Pontential
 - Current Land Use
 - Pontential Land Use
 - Forest Planning
 - Silviculture
 - Harvesting
- (4)Management Practice
 - Ecosystem
 - Systematic
 - Automatical
- (5)Management Results
 - Environment Impact Assessment
 - Cause-Effect Analysis
- (6)Management Adaptive

1. 課程介紹

- (1) 目的
- (2) 內容
- (3) 基本理念

2. 測計學為科學的一支

- (1) 科學認識論與方法論
- (2) 如何找欲鑽研之題目
- (3) 前人研究的查詢—查詢方法、系統、書本與期刊
- (4) 如何研讀科學報告、整理文章
- (5) 如何進行實驗研究
- (6) 如何撰寫科學報告

3. 森林資源經營、管理

- (1) 森林資源的經營與保育觀念(單一目標、多目標、多項資源經營)
- (2) 經營、管理決策支援系統(Decision Support System, DSS)
 - a. 調查、資料收集系統—資料庫(Data Base)
 - b. 資料處理、分析、系統—經營資訊系統(Management Information System, MIS)
 - c. 模式建立、評估系統—模式庫(Model Base)—程式庫(Package Base)
 - d. 資訊提供—資訊庫(Information Base)
- (3) 經營決策系統

4. 經營、管理決策支援系統資料來源

- (1) 來源與資料格式
 - a. 地面調查(數圖、表、像)
 - b. 航空攝影調查(像片)
 - c. 遙感探測(影像、數值)
- (2) 資料間的轉換、整合
 - a. 向量式(Vector)—圖形式(Raster)
 - b. 地理資訊系統(Geographic Information System, GIS)
 - c. 測計、統計、製圖、經濟、資源經營模式

5. 取樣觀念、取樣設計與統計的需要

- (1) 取樣
 - a. 隨機取樣(Random sampling)
 - b. 系統取樣(System sampling)
 - c. 分層取樣(Stratification sampling)
 - d. 雙重取樣(Double sampling)
 - e. 多層級式取樣(Multistage sampling)
 - f. 空間取樣(Spatial sampling)
 - g. 適應取樣(Adaptive sampling)
- (2) 統計
 - a. 描述統計
 - b. 推論統計

- (a)變異數分析
- (b)迴歸分析(線性、非線性迴歸)
- (c)多變值分析
- (3)系統分析、作業研究

6. 電腦科技與測計學

- (1)硬體
 - a.個人電腦(Personal , Professional Computer)、工作站(Workstation)
 - b.周邊設備
 - (a)掃描機(Scanner)
 - (b)數據儀(Digitizer)
 - (c)印表機(Printer)
 - (d)繪圖機(Plotter)
 - (e)攝影記錄器(Recorder)
 - c.網路
- (2)軟體
 - a.文書處理(Word Processing)
 - b.統計分析(Statistical Analysis)
 - c.圖形繪製(Mapping)
 - d.影像分析(Image Processing and Analysis)
 - e.光譜分析(Spectral Analysis)
 - f.空間分析(Spatial Analysis)
 - g.地理資訊系統(Geographical Information System)

7. 測樹學發展與研究方向

- (1)供給
 - a.樹木的組成：根、莖、葉、花、果實、種子
 - b.森林的組成：
 - (a)林木：樹種組成、林分結構、林分生長
 - (b)林地：地理環境、物理環境、化學環境、生物環境
- (2)需求
 - a.樹木的功效：基因、碳儲存、防溫室效應、防污染
 - b.森林的功效：野生動物棲息地、戶外遊憩活動、資源保育、環境美質、水土保持、林木、社會教育
- (3)經營
 - a.林地分類、林地利用方式變更、利用
 - b.林木經營、育林、撫育、保護、伐採
- (4)調查項目、如何調查、調查原理
 - a.除林木胸高直徑、樹高以外的調查項目，如葉、樹冠、根...等
 - b.使用儀器、工具、原理
- (5)資料如何處理、分析
 - 資料格式一致、資料檔之轉換、Chaos、Factorials
- (6)模式如何建立
 - a.處理模式(Process model)

- b.統計模式(Statistical model)
 - c.機制模式(Machanism model)
- (7)資訊如何解釋、如何運用
- 8.大面積森林資源調查、監測的發展與研究動向
 - (1)地面樣區
 - a.永久樣區(Permanent plot)設立
 - b.取樣調查方式的決定
 - (2)航空攝影
 - a.航照片(大小、比例尺、正斜度、重疊度、品質...)的功能與限制
 - b.客觀判釋的研發(經驗與專家)
 - c.土地利用型圖
 - (3)遙感探測
 - a.影像處理
 - b.地理資訊系統的配合使用
 - c.衛星定位系統(Global Sattilite Position System, GPS)配合使用
 - (4)整合三種資料來源之建立與分析、分層、多相取樣分析技術之研發
 - (5)全省資源調查
 - 目標、調查方法、內容、資料格式
 - (6)檢訂調查
 - (7)全省資源調查與檢訂調查的整合
- 9.生長、發展與演替模式之發展與研究動向
 - (1)研究對象
 - a.單株、直徑級、全林分
 - b.樹種、林型
 - c.人工林、天然林(熱帶雨林、溫帶林、寒帶林)
 - (2)研究方法
 - a.測計與生長、經濟
 - b.靜態的形態與現象(空間因子)
 - c.動態的生長與演替(時間因子)
 - d.統計模式、過程模式與機制模式
 - e.模式的誤差與評估
 - (3)生長、演替模式於森林經營、管理上之應用
- 10.森林資源經營管理支援決策系統的動向
 - (1)層級式管理
 - a.分散式管理
 - b.集中式管理
 - (2)網狀式管理

一、農委會遙測計畫

題目：多尺度森林資源調查分析與監測

(Multiscale Forest Resource Inventory & Analysis and Monitoring)

1995.08--1996.07

(一)多尺度森林調查資料整合

- 1.區域性(台灣全省)
空間圖籍與非空間屬性資料空間
不同比例尺圖籍
- 2.不同森林資源層次(台灣全省、惠蓀、埔里、八仙山)整合與查詢系統
樹、樣區、林分、事業區、景觀(林型)、地區等不同尺度資料的整合

1996.08--1997.07

(二)台灣全省森林生態系永久樣區規劃

- 1.建立台灣森林樹種、林型之生態幅度與空間模式
- 2.評估生態系關鍵因子以為永久樣區規劃依據
- 3.由不同森林資源層次、空間，評估永久樣區分布、個數。

1997.08--1998.07

(三)整合森林資源調查與監測

- 1.森林資源調查與生態系監測之比較與整合。
- 2.調查與監測項目、自動化調查技術、方法、過程之評估與確定。
- 3.調查資料管理、資料分析

1998.08--2000.07

(四)(五)森林生態系經營保育決策支援系統

- 1.生物、生態資訊
- 2.社會、經濟資訊
- 3.行政、法規資訊
- 4.經營、保育資訊

二、農委會森林經營管理計畫

題目：森林生態系經營保育決策支援系統

(Forest Ecosystem management and Conservation Decision Making Support System)

1995.08--1996.07

(一)目前經營保育決策問題與原則之探討

- 1.由中央到地方，各層級之經營保育決策問題綜合。
- 2.各行政層級所關注之森林生態系層級項目。
- 3.經營保育決策，所欲依據之資訊。

1996.08--1997.07

(二)檢討已有林業資料庫，規劃森林生態系經營保育所需資料之收集與更新

- 1.檢討已有森林資料庫--林業經營與林學研究
- 2.以森林生態系觀念，經營保育台灣森林，所需之資料與資訊
- 3.不足之資料與資訊之收集、分析規劃

1997.08--1998.07

(三)整合森林生態系經營保育資訊中心

1. 森林生態系經營保育觀念具體化。
2. 適應森林生態系經營之調查系統
含取樣設計、調查技術、資料收集
3. 資料庫管理系統
4. 資料分析與模式建立
5. 決策支援資訊展示。

1998.08--2000.07

(四)(五)森林生態系經營保育決策支援系統

1. 生物、生態資訊
2. 社會、經濟資訊
3. 行政、法規資訊
4. 經營、保育資訊

三、 觀光局東北角風景特定區

題目：陸域生態系資源調查與監測

(Inventory and Monitoring of Terrestrial Ecosystem Resource)

1994.10--1995.11

(一)陸域生態系資源調查與監測(2)

1. 比較GAP\FHMS\LTER，建立生物與生態監測系統。
2. 樣區植生與昆蟲調查設計與實務(取樣與調查方法比較)
3. GIS圖層資料與樣區調查屬性資料建檔、資料管理。
4. 資料分析--植生與昆蟲棲息地關係。

1995.12--1996.11

(二)陸域生態系資源調查與監測(3)

1. 主要土地利用型之植生與鳥類調查設計與實務(取樣與調查方法比較)
2. 資料分析--地形、氣象與土地利用型、植生、鳥類與昆蟲棲息地關係。動植物雜異度之交差分析。
3. 探討本地區環境敏感生物、生態現象。(對遊憩與核能廠建造)

1996.12--1997.11

(三)陸域生態系資源調查與監測(4)

1. 環境敏感生物、生態現象自動監測方法、設備之設計。
2. 監測人員，監測原理、環境敏感生物、生態現象辨認方法、調查方法、步驟、儀器使用之訓練。

1997.12--1998.11

(四)陸域生態系資源調查與監測(5)

1. 整合陸域、水域、海域生態系調查技術與監測手冊
2. 檢討改進自動監測設備

四、國科會計畫

題目：森林族群空間模式之建立與應用

(一) DGPS於森林族群之經營、保育上應用

1. DGPS於森林族群調查上之應用

2. 即時(real time)GPS於森林族群調查上之應用
3. 森林族群空間結構資料庫之建立
4. 整合圖層與資料庫建立

(二) 空間統計於森林生態系經營管理上之發展

1. 以空間統計(或稱地理統計工具,Geostatistical tools)進行空間尺度與結構資料的性態值化和模式化。變異則由共變異變異圖(covariance variograms)或半變異圖(semi-variograms)用來模式化描述地形的空間變異(Dubayah et al., 1989),太陽輻射(Dubayah et al., 1990),表面反照率, albedo(Webster et al., 1989),植生(Pastor & Broschart 1990),光譜植生指標(Davis et al., 1989) 土壤性質(Burrough 1983a, 1983b)。
2. 全省森林生態系空間統計結構的模式之建立與應用
3. 森林族群(林分)空間統計結構的模式之建立與應用
4. 單株林木生長模式之建立與應用

(三) 全林分生長模式、直徑級生長模式與單株林木生長模式之整合

(四) 神經網路、專家系統於生長分析上之研究

環境取樣學 Sampling

高義盛 學號：78833106

第一章 前言

取樣是選取族群的某一部份去觀測，藉由選取某部份去推估整個族群。因此，生物學家在研究區域內選擇地衣的小樣區，去推估可可當馴鹿食物的地衣可利用量。用這些樣本的乾重量為基礎，去推估整個區域的可利用量。相同地，藉由鑽探為數不多（高費用）取樣油井，去推估一個地區可利用油田的數量。民意調查也是相同的情況，只是由人民族群中抽樣，藉由抽樣的意見去比例推估整個人民不同的意見。藉由抽取一些醫療機構的治療病歷，去推估稀少性疾病的流行。在研究區域內，以取樣地點偵測的形式為基礎，去推估稀少性和瀕臨滅絕鳥種的豐富度。

如何去做最佳的取樣和觀測及手邊的取樣資料如何去做最佳的應用去推估整個族群性態值是很明顯的問題。取樣觀測包含樣本個數、如何選取樣本、用什麼觀測方法和如何做量測值記錄。好的觀測推估包括選出適當資料、估計者使用輔助資訊的判斷及推估方法的選擇。

取樣不同於試驗設計，試驗設計可以對干擾族群的行為影響做實驗觀察，取樣是不受干擾去推估族群。因此，問卷調查的用辭不要影響意見的回答，動物族群的觀測不要受動物族群的分布或行為的影響。

取樣不同於單純的觀測研究，它如何獲得族群的觀測值是不可控制。取樣有慎重去選取樣本的機會，因此可以避開事件很多的觀測資料或其他非代表性的不可控制的元素。

1. 取樣和推估的基本觀念

在基本取樣中，族群是已知的、有效的族群 N 個單元——就如同人民或地面樣區。每一單元包含一個有興趣變數的值，就如同這單元的 y 值。在族群中每一單元的 y 值被視為固定值，假如不知道——就不是隨機變數。在族群中的單元被認明和標定為 $1, 2, \dots, N$ 。

族群中只有樣本被選擇和觀察。資料收集涵蓋樣本中每一單元的 y 值一起做單元標定。因此，在油畫中每一油井的鑽探，不僅記錄油量還要明白地對油井位置做標定。除了對有興趣的變數外，任何輔助變動值，就如同油井深度、底質型態都要記錄。在地衣調查中，輔助變數被記錄包括海拔高、其他植被覆蓋率甚至地衣生物量“子實體”的推估。在民意調查問卷中，輔助變數如性別、年齡或收入等級皆會隨意調查被記錄。

從族群中選取樣本單元的程序稱為取樣設計。大多數大家所熟悉的取樣設計，是由每一可能樣本 s 選取樣式的機率 $P(s)$ 表示每一可能樣本 s 。在實際取樣中，取樣設計選取單元的步驟程序被同等地描述，而不是選取整個樣本機率的結果。在簡單隨機取樣中一步一步地選取程序包含從 $[1, 2, \dots, N]$ 隨機選取一個單元標籤，再從 1 到 N 間選取另一個單元標籤，直到 N 個樣本單元被選取。

在基本取樣觀念中族群全部序列 y_1, y_2, \dots, y_N 的 y 值是有一個固定的性態值或族群母數。取樣常用的推理問題是藉由樣本的觀測值去推估一些族群的總性態值，如同 y 值的平均值或總數。此外，大多數的取樣和推估狀況，常常藉由信賴區間的評估去得到推估的正確度或信賴度。

在基本取樣觀念中，藉由取樣的樣本大小去推估族群的 N 單元真值，那麼對有興趣的族群性態值就可以完全地被知道。取樣推估有不確定性是來自於只對族群的部分樣本的觀測。當族群的性態值是固定時，推估依賴的是族群樣本的選取，假如每一可能取樣的推估都相當接近族群性態值的真值時，那麼採取合適取樣策略地不確定性就很小；另一方面假如取樣的不確定性是藉由一個可能的樣本到另一個可能的樣本去推估族群的變異就很大。大多數經常使用的取樣設計策略——最初選用的取樣設計最好——變異性常來自於對單一樣本一再地取樣。

謹慎地選用取樣設計，嚴謹的取樣設計和選用適當的推估方法，能獲得族群總量的推估值的無偏估值。就如同族群平均值或總數的推估是沒有依靠任何族群本身的前題假設。選取相當於真實族群的所有可能的樣本去推估族群期望值是無偏估值的。因此，透過取樣設計和推估程序，地衣生物量的無偏估值推估是由在研究區域內地衣是否全面平均分布或是區塊分布而獲得。除外，逢機或機率取樣選用移除人類已知和未知的偏值來源，就如同在自知或未自知的意向下去取樣比對有興趣變數的平均值較多或較少的單元。特別值得注意的是取樣調結果並不可能符合所有對此一調查有興趣的人的需求。例如魚類族群調查結果並不能同時滿足漁業管理者、漁販、環境保護論者的需求。在所有事件中，族群代表性取樣的選擇目標未必被全部團體所認同。

機率設計就如同簡單逢機取樣，藉由評估調查結果的可信賴度去提供族群平均值或總數及變異性的無偏估值的推估。無偏估值推估和變異性推估可由不等機率設計獲得，藉由已知的每一單元和雙單元的所有機率樣本。

取樣調查的無偏估或接近無偏估的推估目的，是反應調查結果的精確度或變異性的指標，並能使取樣調查便利或低成本。在環境多變化下，取樣設計和推估方法廣泛的發展，以盡可能滿足這些目標的需求，包括簡單逢機和不等機率取樣，輔助資訊的使用，分層、系統、多層次和雙重取樣，其他取樣技術。

2. 取樣單元

人民和團體的很多大族群，對實際取樣的觀測問題，先分明取樣的單元形式和敘述單元表列的架構。單元也許是人民、家庭、醫院或商店。在目標族群中，人民、醫療機構或商店的完全表列是藉由理想架構取樣單元選擇而獲得。實際上，獲得真正符合有興趣族群的表列常常是有困難的。電話簿不能表列沒有電話或沒有登刊的家庭，所有可能有電話的家庭也許可以做逢機電話取樣，但不能包括沒有電話的家庭，且公家私人的電話簿並非即時更新。

有很多其他的族群並沒有清晰明瞭的單元，在一個區域的自然資源調查或農作物調查，也許被區分成地理單元（樣區或部分），不管如何任意去選擇不同的單元大小和形狀，將會影響調查的成本和推估精確度。因此，取樣的程序是在

研究區域內隨機選擇點的位置，取樣單元要能涵蓋過去不是有限的被選擇樣本的族群單元。

在無法辨明族群的偵測問題上，取樣單元和樣區的規則是藉由族群觀測的方法和被選擇觀測的位置的偵測函數來使用。例如，鳥類調查線截法的位置和飛行速度的選擇，即可決定研究區域內“有效地區”觀測的傳統取樣單元和樣區。

有些取樣的情況，有興趣的變數變化是一個區域連接一個區域（連續性區域變化）。例如，一個地區的油田評估調查，變數的測量可能是藉由測量一個位置的深度或核心容量。在區域內變數的值並不需要和任何有限組成單元相結合，但可藉由亞區的大小和形狀，來測量或推估樣點或整個區域。

雖然，前述的取樣狀況超越一個區隔成唯一有限收集的被取樣樣本的單元，基本取樣設計考慮的是有關隨機取樣、分層取樣和其他設計及設計一不偏估值的推估、比例推估和其他的方法。

3．取樣和非取樣誤差

基本取樣的觀念認為對有興趣的變數每一單元的取樣測量是沒有誤差的，推估誤差只包含取樣的部分族群，因此，誤差被當“取樣誤差”提及。但是真正的測量情況，非取樣誤差也會發生。在電話問卷中，取樣時因某些人不在家而不被選取或拒絕回答問卷中的問題，不回答問卷者並不具整個族群的代表性，如此族群的取樣傾向於非代表性且推估值是具偏差的。在魚類調查中，由於惡劣的氣候狀況，有些地區不被選擇觀測；地點由沿岸到遠洋，研究區域因氣候問題而不能代表整體魚類族群。

沒有回答的問題，特別是在調查中宣布是非常低回答率的，其回答機率與其性態值的量測有關——尤其在雜誌讀者性別調查上的實證問題。沒有回答問題的影響也許藉由增加取樣的努力來推估族群沒有回答層的性態值，明智的利用回答與沒有回答單元可利用的輔助資訊或沒有回答情況下的模式，但是最佳的忠告是力求沒有回答率盡可能的降低。

誤差在有興趣的變數的測量和記錄中都有可能發生。質的控制是藉由調查的每一階段都要保持最小誤差的努力。有些狀況，模式量測誤差個別地由有關族群性態值觀測的取樣結果發生。

偵測問題非取樣誤差的形態是發生在無法辨明族群的廣泛區域。鳥類調查中，在被取樣地點的附近，觀測者一般是難以發覺鳥種的每一個體。以拖網做魚類調查，並不是每一種魚都會被魚網所補捉。在做社會調查並沒有包含每一個無家可歸的人。物種調查技術包括線截法、補捉—再補捉

和其他已經被發展的以偵測為中心的推估的相關方法。

4．取樣模式

在基本取樣觀念中，族群是有限的單元組合，每一單元有一有興趣的變數的固定值樣本單元被選取的程序是經由機率來設計。但是有些族群，實際利用族群本身的機率模式，模式是基於族群的形態分布影響的自然現象的知識或實際的統計模式包含族群的基本性態值。

例如，迴歸模式實證描述有興趣的變數間的關係，園藝作物的收益藉由輔助變數來判定，就如同空氣污染以中位數標準來判定。模式有興趣的變數與輔助變數間的相關模式，包含如何去設計調查和如何進行推估二者。

在空間取樣的情形，不同地點有興趣的變數的值間的相關是存在的，依靠地點間的距離，包括關於取樣設計的選擇，推估、預測和觀測的芳法。一個模式建立特別是在礦物和化石燃料資源地地質調查是基於利用相關形態的影響。在生態調查上，相關形態被定義不單是關觀測地點的選擇，還包括觀測方法（包括樣區形狀）的使用。

在理論上，利用一個模式來解釋族群，但是並非所有的調查都要依靠模式來解釋。粗放的取樣調查使用模式來顯示有效的而避免偏離模式的設計程序。

5. 適應性和非適應性設計

稀少性、群團族群的調查動機是比取樣基本觀念還要進步。在適應性取樣設計，程序依賴有興趣的變數的觀測記錄去選擇地點或單元去觀測。例如，自然資源豐富度推估的調查，調查期間在高豐富度觀測地點附近增加地點去增加取樣。如此適應性設計在動物、植物、礦物、化石燃料資源和其他領域如遺傳學、品質管制都是很重要的應用。

適應性設計程序地主要目的是利用族群性態值觀測去比較相同樣本大小的傳統設計，以獲得取樣的精確度或有效性。適應性設計的程序包含連續性中止的規則和層間連續性配置的程序——在有限族群情況下做連續性分析的程序是艱辛的研究。取樣的狀況下族群單元被認定，在適應性設計程序的機率是較大的，因為可能去決定調查期間不只是如何從很多單元去做取樣，還要考慮下一個取樣的單元或集團單元。

在適應群團取樣，對有興趣的變數的值的觀測必須滿足的原則——例如，在一個地點觀測動物高豐富度——觀測地點的鄰近地區必須增加單元去取樣，在本書最後一章會對不同變異方——法做描述。有些族群適應性設計的結果，對於稀少性、群團族群有明顯增加其效率和接近現狀。

取樣設計藉由選取特別樣本 s 的機率 $P(s)$ 的取樣來做傳統或非適應性設計。適應性設計選取樣本單元的機率是 $P(s|y)$ ，選取樣本 s 的條件，就是在有興趣的變數的族群中 y 值組成中。當然，實際上選取過程不單是依靠那些值，還包含已觀測值的部份。

很多自然族群有集中於研究區域中的小區域的趨勢，但是他們集中的位置並不能在調查前先做預測。對族群做有效的適應性設計能使那些集中區域優勢單元的高選取機率取樣的結果。適應性設計的首要目的是去獲得族群總數較精確的推估，第二效益是對有興趣觀測的部分做戲劇性的取樣增加——例如，較多動物被看見或較多礦物被獲得。因為適應性設計被考慮到，取樣方法學的範圍和潛在力將會非常寬廣。

第八章 迴歸推估 (Regression Estimation)

在有些取樣的情況，有興趣 y 的變數大概與輔助變數 x 是線性的關係，但當 $x=0$ 時， $y \neq 0$ 。例如，假使植物作用每單位面積產量 $=y$ ，那樣區附近的空中污染平均集中量 $=x$ ，當 y 因 x 而增加，當 $x=0$ 時， y 增加至最大。有些情況線性迴歸推估比比比例推估還要適當好用。

實際上，更多輔助變數與每一單位有關。前述取樣，輔助變數 x_1 也許是大氣中二氧化硫的平均值，輔助變數 x_2 也許是大氣中氧化氮的平均值，另一輔助變數 x_3 也許是一個適當指示變數，假如對土壤型最適與否加以分類的。複迴歸分析模式或一般線性統計模式，可以敘述有興趣變數和許多輔助變數間的關係。

最初，藉由一個輔助變數的線性迴歸推估的被描述，在基礎設計或固定族群的狀況。不等機率迴歸推估設計和複迴歸模式，在這一章最後一節描述。像比例推估、迴歸推估是在簡單逢機取樣下的非無偏估值設計。通常在迴歸模式下的假設，推估是無偏估值的。

假使迴歸模式描述輔助變數與有興趣的變數間的猜測關係是可行的，那取樣的主要目的是對族群的某些性態值 y 值的預測。性態值的預測也許是取樣中族群平均數或總數或單一單元的 y 值與否。簡單而言，線性預測的基本結果，接近於一個輔助變數的簡單線性迴歸模式及多個輔助變數的複迴歸模式。

1. 線性迴歸推估

族群中第 i 單元的推測是從 $i=1, 2, \dots, N$ 輔助變數的 x_i 值的和有興趣的變數的 y_i 值的結合，族群平均數和總數的 y 值分別以 μ 和 τ 表示，族群平均數和總數的 x 值分別以 μ_x 和 τ_x 表示，簡單逢機取樣的 n 單元被選取。線性迴歸推估的族

群平均數 μ 的表示 $\hat{\mu}_L = a + b\mu_x$ 在 $b = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / \sum_{i=1}^n (x_i - \bar{x})^2$ 和 $a = \bar{y} - b\bar{x}$ 。

資料藉由最小平方方法求得適當直線的 a 值是 y 的截距和 b 值是斜率。 a 可以取代，推估者可寫成 $\bar{\mu}_L = \bar{y} + b(\mu_x - \bar{x})$ 。

像比例推估，線性迴歸推估在簡單逢機取樣是非無偏估值的設計。

1981年Scott and Wu在有限族群中央極限定理對於迴歸推估市基於信賴區間的步驟上。然而，Royall和Cumberland(1985)對真實族群樣本大小 $n=32$ 的迴歸和比例推估的實驗研究顯示真正所包含的機率用標準變異數推估就如同計畫中二擇其一，也許較名義上的信賴水準還要低。Deng和Wu(1987)在隨後理論上的研究比較相互變異數的推估和假定不同的推估者所使用。假定推

估平均值平方誤差或使用信賴區間的概念，這方面的課題是需要更多的研究。

2. 不等機率迴歸設計

Horvitz-Thompson推估是取樣設計包括第*i*單元的機率是 π_i ， $i=1, 2, \dots, N$ ，族群平均數 μ 的y值無偏估值推估。 $\hat{\mu}_y = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\pi_i}$ 在總和超過取樣的 ν 明顯單元，輔助變數x在Horvitz-Thompson推估是基於樣本x值。 $\hat{\mu}_x = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\pi_i}$ 族群平均值的x值是無偏估值推估。

在一般迴歸推估，族群平均值在近乎無偏估值下推估 $\hat{\mu}_G = \hat{\mu}_y + \hat{B}(\mu_x - \hat{\mu}_x)$ \hat{B} 是加權迴歸斜率推估，在包含機率的基礎上。

3. 迴歸模式

像比例推估，迴歸推估在簡單逢機取樣設計下，是非無偏估值的，就是y和x值是固定值的觀點，期望值超過所有可能的取樣，族群平均值的y值的迴歸推估不等於真正族群的平均值。在一個正常線性迴歸模式的假設，其迴歸推估是無偏估值的。

假定一個標準線性迴歸模式是成立的。那 Y_1, \dots, Y_N 是依賴x值分布的逢機變數。 $E(Y_i) = \alpha + \beta x_i$ $i=1, 2, \dots, N$ 。不同單元的Y值被假設為不相關且變異數固定。 $\text{var}(Y_i) = \sigma_L^2$ 對所有i變異數 σ_L^2 被認為是真正迴歸線的Y值變異，藉由參數 α 和 β 決定。

因為族群Y值是逢機變數，他們的任何的函數就如他們的平均值 $\mu = \left(\frac{1}{N}\right) \sum_{i=1}^N Y_i$ 是逢機變數。我們希望藉由觀察取樣的值去預測逢機變數的值 μ ，並能顯示(特別在Royall的報告)在最小平均平方誤差下的模式的線性預測是線性迴歸推估的 $\hat{\mu}_L$ 。在模式下，線性迴歸推估是模式不偏估值，就是 $E(\hat{\mu}_L) = E(\mu)$ 在期望值模式下的假定，樣本S的無偏估值是成立的。在樣本S中，平均值平方預測誤差模式是

$$E(\hat{\mu}_L - \mu)^2 = \frac{N-n}{N} \left[\frac{1}{n} + \frac{N}{N-n} \frac{(\bar{x} - \mu_x)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \sigma_L^2 \quad \text{平均值均方誤差的無}$$

偏估值的獲得，藉由以 σ_L^2 置換上式，樣本誤差變異數

$$\hat{\sigma}_L^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - a - bx_i)^2 \quad \text{然而較簡單的推估(等式1)，被粗放的}$$

考慮由假定模式中分離。

迴歸模式和預測方法在調查取樣中被發展和討論在Cumberland和Royall(1981, 1988)....。從調查資料推估假定模式迴歸參數的問題被討論在Holt和Scott(1981)....。

4 · 複迴歸模式

迴歸模式容易地由一個輔助變數擴大至多個變數，讓 X_{i1} , X_{i2} , ..., X_{ip} , 藉由第i單元與P輔助變數結合，這就是所謂的複迴歸模式。 $E(Y_i) = \sum_{K=1}^P \beta_K X_{iK}$ $\text{cov}(Y_i, Y_j) = \sigma_L^2 \nu_{ij}$ $i=1, \dots, N$ $j=1, \dots, N$ 簡單線性迴歸和比例模可獲得這些特別的例子。

假定一個希望去預測一些線性組合族群y值的 $\sum_{i=1}^N l_i Y_i$ ，例如每一個係數 $l_i = 1$ ，有興趣的函數是族群總數，每一 $l_i = 1/N$ ，它是平均值。假如希望單單去預測一個非取樣單元的值，假定單元j那其他N-1單元的 $l_j = 1$ ， $l_i = 0$ 。

線性預測問題上發現一個線性函數 $\sum_{i=1}^n a_i Y_i$ 的觀測值是族群數量的無偏估值且是最小可能均方差預測，可發現 a_1, a_2, \dots, a_n 值

，可用 $E\left(\sum_{i=1}^n a_i Y_i\right) = E\left(\sum_{i=1}^N l_i Y_i\right)$ 表示。最小均方差預測以

$$E\left(\sum_{i=1}^n a_i Y_i - \sum_{i=1}^N l_i Y_i\right)^2 \text{ 表示。}$$

母體標記模式是 $E(Y) = X\beta$ $\text{var}(Y) = \sigma_L^2 V$ n個單元的樣本S，Y向量可以定義成最初n位置放置取樣的Y值而且可以寫成劃分形式 $Y = \begin{pmatrix} Y_s \\ Y_r \end{pmatrix}$ Y_s 是取樣值n*1的向量， Y_r 是在其餘族群(N-n)*1向

量的值。相同地， $N \times P$ 次元母數 X 可以劃分為 $X = \begin{pmatrix} X_s \\ X_r \end{pmatrix}$ 相同地，

變異數-協方差母數 V 可以劃分為 $V = \begin{pmatrix} V_{ss} & V_{sr} \\ V_{rs} & V_{rr} \end{pmatrix}$

預測的基礎的取樣可寫成 $a'Y_s$ ，當向量係數 $a = (a_1, \dots, a_n)'$ ，預測的數值可以寫成 $I'Y = I'_s Y_s + I'_r Y_r$ 當 $I = (I'_s, I'_r)$ 且 I_s 和 I_r 包含取樣單位的 1 值和代表其他族群。

5. 迴歸模式的設計應用

從等式 2，假定取樣的線性迴歸模式的平均值均方誤差是最小的，假設取樣被選取的 $\bar{x} = \mu_x$ ，也就是 x 是平衡的。當取樣的 x 是平衡的，迴歸推估等於樣本 y 值的平均值，所以平衡的 x 對假定迴歸模式是粗放推估。在 Royall 和 Cumberland (1981b) 和 Cumberland 和 Royall (1988) 的研究更深一層顯示，當平衡樣本被選取時，均方誤差對正常推估最佳。

更多一般線性模式，可能決定取樣的最小等式 3，從假定模式對抗粗放取樣的分離，每一 x 變數的取樣平衡也就是取樣的 $\bar{x}_k = \mu_k$ ， $k = 1, \dots, p_1$ 是被要求的。不等機率取樣，一個代表性取樣被定義為 $\mu_k^A = \mu_k$ 對於每一 P 輔助變數，就像一個取樣一般迴歸推估以 y 值為基礎的 Horvitz-Thompson 推估是相同的。

除此，迴歸模式 $E(Y_i) = \sum_{K=1}^P \beta_K X_{iK}$ ， $\text{var}(Y_i) = \sigma^2 v_i$ ， Y 是獨立的。每一包括機率 π_i 的比例項設計對 v_i 平方根是有效的。在模式組合下均方差是最小的且設計的一般迴歸推估近乎無偏估值。設計遠選取產生平衡樣本的步驟，在相同時間個別期望樣本的選取機率也許不是每一次的任務。分層逢機取樣藉由輔助變數分層的基礎和層內 v_i 平均值分布的基礎可能是密切的。

假設擇取在不是任何逢機狀況下是錯誤的，當不知其他一個或多個輔助變數的重要性時，取樣藉由輔助被平衡的調查例子，在一個研究區域去推估甲蟲的豐富度。甲蟲的豐富度被認為與海拔高 (x) 近乎線性相關，所以取樣樣區被有系統選取在研究區域內。在低海拔到高海拔的路旁設置，所以 $\bar{x} = \mu_x$ 。研究者不得而知地，在研究地區真正影響甲蟲豐富度的重要變數是 X_2 ， X_2 是離道路的距離和其它生育地的干擾。因此，對平衡取樣是全不具代表性的且導出錯誤的推估結果。當然，研究者能重覆調查在已知輔助變數間的平衡嗎？假如沒有更深一層探討是不知道重要的輔助變數。很多理由原因，很多研究者令人安慰的是藉由相同基準的複合，逢機取樣設計—儘量已知輔助變數的分層基礎仍就被使用。

第十二章 群團和系統取樣(Cluster and Systematic Sampling)

雖然系統取樣和群團取樣表面上看是相反的-前者是樣本單元的區隔，而後者是樣本群集在一起-兩種取樣設計使用相同的樣本。族群分割成第一單元，而每一個第一單元是由第二單元所構成。而第一單元被包含在樣本中，在樣本中之每一第二單元被觀察。

在系統取樣中，有些遍及族群的系統性形成的包含第二單元的第一單元被分離。在群團取樣中，包含第二單元的第一單元經常彼此間相接近。在空間組成中系統取樣第一單元是由研究區域網格形式的地面樣點所組成。群團第一單元包括空間排列如鄰近樣點四周或長、窄條的鄰近單元。在圖12.1顯示群團取樣包含十個簡單逢機取樣的第一單元，每一個第一單元包含八個第二單元。圖12.2顯示系統取樣有兩個逢機選取的起始點。系統取樣包含兩個第一單元(以陰影區分)，每一個第一單元包含16個第二單元。

在任何系統或群團排列的關鍵點，一個第一單元的任一第二單元被包含在樣本中，這個第一單元的所有第二單元都被包括。有些真實的測量選取第二單元是透過第一單元的選取。原則上，關於第二單元的分配是根據第一單元如取樣單位和利用，像任一第一單元有興趣的變數包含第二單元 y 值的總數。所有推估的性質獲得是基於樣本第一單元被選取的取樣設計。然而，幾項系統取樣和群團取樣的共同特點在取樣設計時值得考慮像特殊的項目：1. 在系統取樣中，樣本大小是共通性，就是單一的第一單元。2. 在群團取樣群團的大小是有用的，像輔助資訊被使用來選取不等機率的群團或形成比例推估。3. 群團的大小和形狀會影響取樣效率。

以 N 表示族群第一單元的數量，以 n 表示樣本第一單元的數量。以 y_i 表示第 i 個第一單元的第二單元的數量。在族群中第二單元的總數 τ 。以 y_{ij} 表示第 i 個第一單位中第 j 個第二單元有興趣變數的值， y_i 表示第 i 個第一單位的 y 值總數。族群總數 τ 。每一第一單元的族群平均數 \bar{y}_i 。每一第二單元的族群平均數 \bar{y} 。

1. 簡單逢機取樣選取第一單元

無偏估值推估

當第一單元藉由不置換簡單逢機取樣被選取，族群總數 τ 的不偏估值推估是

$$\hat{\tau} = \frac{N}{n} \sum_{i=1}^n y_i = N\bar{y} \dots (1) \quad \text{樣本第一單元總數的平均值 } \bar{y} = \left(\frac{1}{n}\right) \sum_{i=1}^n y_i \quad \text{。變異數}$$

的推估 $\text{var}(\hat{\tau}) = N(N-n) \frac{\sigma_\mu^2}{n} \dots (2)$ ，有限族群第一單元總數的變異數

$$\sigma_\mu^2 = \frac{1}{N-n} \sum_{i=1}^N (y_i - \mu_1)^2 \quad \text{，變異數 } \hat{\tau} \text{ 的無偏估值推估 } \text{var}(\hat{\tau}) = N(N-n) \frac{s_\mu^2}{n} \quad \text{，樣本}$$

第一單元總數的變異數 $s_\mu^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$ 。這些結果類似簡單逢機取樣。每

一第一單元 μ_1 平均值的無偏估值推估是 $\bar{y} = \hat{\tau}/N$ ，每一第二單元 μ 平均值的無偏估值推估是 $\hat{\mu} = \hat{\tau}/M$ 。 \bar{y} 的變異數是 $\text{var}(\bar{y}) = (1/N^2)\text{var}(\hat{\tau})$ ， $\hat{\mu}$ 的變異數是 $\text{var}(\hat{\mu}) = (1/M^2)\text{var}(\hat{\tau})$ 。推估變異數類似藉由區分 N^2 或 M^2 的 $\hat{\tau}$ 的推估變異數獲得。

比例推估

假如第一單元總數 y_i 與第一單元大小 M_i 是高度相關，基於樣本大小的比例推估也許是有效的。族群總數的比例推估是 $\hat{\tau}_r = rM$ ，樣本比例 r 是

$$r = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n M_i}.$$

族群比例是每一第二單元 μ 的平均值，當比例推估， $\hat{\tau}_r$ 不是無偏估值，但偏估值在大樣本大小是趨向於很小，均方差也許較少考慮，當 y_i 和 M_i 趨向於比例相關的無偏估值推估。

$$\text{均方差或比例推估的變異數的近似公式 } \text{var}(\hat{\tau}_r) \approx \frac{N(N-n)}{n(N-1)} \sum_{i=1}^N (y_i - M_i \mu)^2,$$

$$\text{變異數的推估 } \text{var}(\hat{\tau}_r) = \frac{N(N-n)}{n(n-1)} \sum_{i=1}^n (y_i - rM_i)^2 \text{ 或調整過的推估}$$

$$\text{var}(\hat{\tau}_r) = \left(\frac{nM}{N \sum_{i=1}^n M_i} \right)^2 \text{var}(\hat{\tau}_r).$$

推估每一第一單元 μ_1 的族群平均，比例推估是 $\hat{\mu}_r = \hat{\tau}_r/N$ ，藉由均方差公式 N^2 來區分。推估每一第二單元 μ 的族群平均，比例推估是 $\hat{\mu}_r = \hat{\tau}_r/M = r$ ，藉由均方差公式 M^2 來區分。

2. 取樣機率隨樣本大小成比例來選取第一單元

假定第一單元的選取是經由第一單元大小成比例的機率以描繪置換取樣來選取 $p_i = M_i/M$ 。由族群 M 中選取 n 個第二單元取樣設計，是採用可置換的簡單逢機取樣：任一第二單元的選取是經由第一單元的選取。

Hansen-Hurwitz(PPS)推估

族群總數無偏估值推估是基於Hansen-Hurwitz(PPS)推估在取樣機率隨樣本大小成比例的可置換取樣下進行。 $\hat{\tau}_p = \frac{M}{N} \sum_{i=1}^n \frac{y_i}{M_i}$ 變異數推估

$$\text{var}(\hat{\tau}_p) = \frac{M}{n} \sum_{i=1}^N M_i (\bar{y}_i - \mu)^2 \quad \bar{y}_i = y_i/M_i \quad \text{變異數無偏估值推估}$$

$\text{var}(\hat{\tau}_p) = \frac{M^2}{n(n-1)} \sum_{i=1}^n (\bar{y}_i - \hat{\mu}_p)^2$ $\hat{\mu}_p = \hat{\tau}_p/M$ 。推估 $\hat{\mu}_p$ 是每一第二單元 μ 族群平均的無偏值在取樣機率隨樣本大小成比例的狀況下選取， $\hat{\mu}_{1p} = \hat{\tau}_p/N$ 是每一第一單元族群平均的無偏值。推估變異數的公式是藉由 M^2 或 N^2 來表示變異數 $\hat{\mu}_p$ 。

Horvitz-Thompson推估

Horvitz-Thompson推估是使用機率的取樣設計來估算 $\pi_i = 1 - (1 - p_i)^2$ 複

合取樣機率 $\pi_{ij} = \pi_i + \pi_j - [1 - (1 - p_i - p_j)^n]$ 是基於選取機率 $p_i = M_i/M$ 。族群總數的Horvitz-Thompson推估是 $\hat{\tau}_\pi = \sum_{i=1}^v \frac{y_i}{\pi_i}$ v 樣本第一單元的數量。推估變異

數的公式 $\text{var}(\hat{\tau}_\pi) = \sum_{i=1}^N \left(\frac{1 - \pi_i}{\pi_i} \right) y_i^2 + \sum_{i=1}^N \sum_{j \neq i} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} \right) y_i y_j$

3. 基本原則

因為每一第二單元的觀測是在被選取的第一單元內，而第一單元內的變異數並不能進行變異數推估。因此，基本系統和群團取樣的原則是去獲得低變異數或均方差的推估，族群被分成以相似的群集成為群團，同樣地，第一單元內的變異數盡可能的大，以獲得族群或總數的最大精確推估。理想的第一單元要能包括族群的全部歧異度，才有代表性。

植物、動物或礦物及人類族群等自然族群的空間分布，是典型的滿足系統取樣的第一單元的典型，而第二單元是被分開的，但不包括地理鄰近單元的群團。群團取樣常常不能實行的便利或實用的理由是由第二單元觀測的數量去獲取最低的變異數。

4. 單一系統取樣

很多調查使用系統取樣設計來逢機選取單一的起始單元，進而適當分隔來觀察次一單元。這種取樣包含逢機選取的單一第一單元。從樣本的大小，可能獲得族群平均數或變異數的無偏估值，但不可能或得樣本變異數的無偏估值。

無庸置疑，在單一系統取樣第一單元後之第 M_1 次一單元是從族群中 M 個次一單元以簡單逢機選取，而且如果族群具有逢機規則時可以使用簡單逢機取樣變異數的公式去做較佳的變異數推估。有很多自然族群，鄰近族群彼此間很相似，這個取樣過程傾向於對族群平均數或總數的變異數推估有高估現象。

Bellhouse(1988)、Murthy and Rao(1988)、Wolter(1984)對簡單系統取樣變異數推估的變化有所討論。如果分層逢機取樣被使用，對鄰近單元的族群組合成層並做變異數推估是最簡單的方法之一

例12.1 在禿鷹巢的調查，1300公里的海岸線被劃分為100公里長的單元，所以族群單元數 $M=13$ 。系統取樣逢機選取起始三單元中的一個單元，然後每隔第三個單元被選取。因為樣本包含每一單元，使用研究船和飛機來取樣調查所有的禿鷹巢，因此使用簡單逢機取樣選取，不同可能的樣本大小 M_i 。第一單元數目是 $N=3$ ，有 $n=1$ 被選取。假如最初單元被選取當起始點，那 $M_1 = 5$ 將被觀測，然而另一起始點被選擇， $M_i = 4$ 將被觀測。

假定由最初三個單元中逢機選取單元3，在調查中被觀測的 y 值是5, 1, 10 and 18個巢數，那海岸線禿鷹巢總數無偏估值的推估是 $\hat{\tau} = \frac{3}{1}(5+1+10+18) = 102$ 。

然而，從調查的資料中，很多人選擇使用四個觀測值的平均數來推估族群數目

$\hat{\tau}_r = \frac{5+1+10+18}{4}(13) = 110.5$ 在比例推估中，這種推估不是無偏估值。

由簡單起始點($n=1$)，取樣機率隨著樣本大小成比例(PPS)的推估被確認在比例推估。因此，假如系統取樣是與取樣機率隨著樣本大小成比例(PPS)，那上述第二種推估方法是無偏估值推估。這將完成從族群的13個單元中逢機選取樣本，樣本單元是左右每隔的第三單元。

5. 群團和系統取樣的變異數和費用

群團和系統取樣的效益依賴由第一單元選取的大小和形狀的變異數和取樣單元數目的費用成本。當一個起始點，選取 n 個第一單元的變異數與簡單逢機取樣相等的第二單元做比較。群團的平均大小 $\bar{M} = M/N$ ，所以在簡單逢機取樣的 n 個第一單元的第二單元的期望數目 $n\bar{M}$ 。

族群總數的無偏估值推估是基於簡單逢機取樣的 $n\bar{M}$ 第二單元 $\hat{\tau}_{srs} = M\bar{y}$ 這種設計推估結合的變異數 $\text{var}(\hat{\tau}_{srs}) = M^2 \frac{(N\bar{M} - n\bar{M})}{nN\bar{M}^2} \sigma^2 = N^2 \left[\frac{\bar{M}(N-n)}{nN} \right] \sigma^2$

是第二單元的有限族群的變異數 $\sigma^2 = \sum_{i=1}^N \sum_{j=1}^{\bar{M}} (y_{ij} - \mu)^2 / (N\bar{M} - 1)$ $\mu = \tau / (N\bar{M})$ 。

群團或重覆系統取樣，簡單逢機取樣的 n 個第一單元，無偏估值以 $\bar{\tau}_\mu$ 表示， μ 表示使用逢機取樣第一單元的類型 μ 的推估設計。 μ 表示第一單元的大小、形狀和排列，例如正方形群團、長方形群團或系統性樣本。這種設計推估結合的變異數 $\text{var}(\hat{\tau}_\mu) = N^2 \left(\frac{N-n}{nN} \right) \sigma_\mu^2$ $\sigma_\mu^2 = \sum_{i=1}^N (y_i - \mu_1)^2 / (N-1)$ $\mu_1 = \tau / N$ 。

對於簡單逢機取樣的相等樣本大小的群團或系統取樣的相關效益，被定義為變異數割比例 $\frac{\text{var}(\hat{\tau}_{srs})}{\text{var}(\hat{\tau}_\mu)} = \frac{\bar{M}\sigma^2}{\sigma_\mu^2}$ 假如變異數 σ_μ^2 是介於第一單元與全部族群變異數 σ^2 間群團或系統取樣是有效的。

由群團或系統取樣設計去推估使用資料的相對有效性，有用的取樣變異數 s^2 不能使用 σ^2 來推估，因為資料不是由簡單逢機取樣獲得。替代的， σ^2 能推估群團或系統取樣的變異數分析。

假定 N 個第一單元的每一個有相同的第二單元 \bar{M} ，族群的總和平方

$\sum_{i=1}^N \sum_{j=1}^{\bar{M}} (y_{ij} - \mu)^2 = \sum_{i=1}^N \sum_{j=1}^{\bar{M}} (y_{ij} - \bar{y}_i)^2 + \bar{M} \sum_{i=1}^N (\bar{y}_i - \mu)^2$ $\bar{y}_i = \sum_{j=1}^{\bar{M}} y_{ij} / \bar{M}$ 第一項在右邊部是第一單元內的總和平方及第二項是第一單元之間的總和平方。

簡單逢機群團取樣的 σ^2 無偏估值推估 $\hat{\sigma}^2 = \frac{N(\bar{M}-1)s^2 + (N-1)\bar{M}s_b^2}{N\bar{M}-1}$ 群團取樣推估的相對效率是基於來自群團取樣的資料 $\hat{\sigma}^2 / \bar{M}s_b^2 = \bar{M}\hat{\sigma}^2 / s_\mu^2$ 。用相同大小群團的群團取樣能比較較小單元的變異效率。

群團或系統取樣的變異數能由第一單元內的相關條件中檢驗。第一單元的相關係數被定義為 $\rho = \frac{\sum_{i=1}^N \sum_{j=1}^{\bar{M}} \sum_{j' \neq j} (y_{ij} - \mu)(y_{ij'} - \mu)}{(\bar{M}-1)(N\bar{M}-1)\sigma^2}$ 第一單元變異數 σ_μ^2 的總和

平方 $\sum_{i=1}^N (y_i - \mu_1)^2 = \sum_{i=1}^N \sum_{j=1}^{\bar{M}} (y_{ij} - \mu)^2 + \sum_{i=1}^N \sum_{j=1}^{\bar{M}} \sum_{j' \neq j} (y_{ij} - \mu)(y_{ij'} - \mu)$ 群團取樣的變異數 $\text{var}(\hat{\tau}) = \frac{N^2(N-n)}{nN} \left(\frac{M-1}{N-1} \right) \sigma^2 [1 + (\bar{M}-1)\rho] \approx \frac{N^2\bar{M}(N-n)}{nN} \sigma^2 [1 + (\bar{M}-1)\rho]$

假如 $\rho=0$ 群團取樣的變異數和第二單元相同數量的簡單逢機取樣的變異數幾乎相同，假如 ρ 大於0 簡單逢機取樣將有較低的變異數，假如 ρ 小於0 群團取樣將有較低的變異數。

很多自然族群單元彼此間相類似，所以組成群團 ρ 是大於0。有些族群，第二單元關係密切的正方形群團的 ρ 值及 \hat{t} 的變異數較第二單元彼此間疏離的長方形、窄形群團趨向於較大。系統取樣由於每一第一單元的第二單元彼此間是分離的，所以 ρ 可能小於0。由於這個理由，系統取樣與很多真實族群較相符合。

原則上，第一單元理想的大小和形狀能決定每一依賴第一單元大小和形狀變異數和成本函數。在真實的取樣狀況函數不是簡單單純的。Cochran(1977)、(1978)、Kish(1965)和Hansen、Hurwitz和Madow(1953)。

第十三章 多層級取樣設計(Multistage Designs)

假如選取第一單元樣本後，從每一選取的第一單元中選取第二單元的樣本，這種取樣設計稱為二階段取樣。假如從每一選取的第二單元中選取第三單元的樣本，這種取樣設計稱為三階段取樣。更多層取樣設計都是可能的。

多層級取樣設計被使用於實際狀況的變異。商業漁業的漁獲量取樣，必須先對船舶做取樣，再對每一被選取的船舶做漁獲量的取樣。植物物種的取樣，以先做地面樣點取樣較便利，然後再對被選取的地面樣點做植物取樣。

單一階段的取樣設計像群團取樣，因為第一單元不同樣本會有不同的推估變異性發生。二階段取樣設計，第一單元選取的變異推估，因為第二單元不同的次樣本將提升推估的不同值。

N 表示族群第一單元的數量， M_i 表示第 i 個第一單元的第二單元的數量。 y_{ij} 表示第 i 個第一單元的第 j 個第二單元的有興趣變數的值。第 i 個第一單元 y 值的總和 $\sum_{j=1}^{M_i} y_{ij}$ ，第 i 個第一單元的每一第二單元的平均 $\bar{y}_i = (\sum_{j=1}^{M_i} y_{ij}) / M_i$ ，族群總和 $\sum_{i=1}^N \sum_{j=1}^{M_i} y_{ij}$ ，每一第一單元的族群平均 $\bar{y} = (\sum_{i=1}^N \sum_{j=1}^{M_i} y_{ij}) / N$ ，每一第二單元的族群平均 \bar{y}_i ， $\sum_{i=1}^N M_i$ 是族群第二單元的總數。

圖13.1顯示二階段取樣設計，第一階段簡單隨機取樣選取10個第一單元，第二階段由每一被選取的第一單元中 M_i 的第二單元簡單隨機取樣選取 m_i 個。圖13.2顯示第二單元相同總數的二階段取樣設計，但選取 n 個第一單元，每一被選取第一單元中選取 m_i 個第二單元。

1. 每一取樣階段都是簡單隨機取樣

無偏估值推估

簡單隨機取樣使用第二階段取樣，取樣第 i 個第一單元的 y 值總數的無偏估值推估 $\hat{y}_i = \frac{M_i}{m_i} \sum_{j=1}^{m_i} y_{ij} = M_i \bar{y}_i$ (公式1) $\bar{y}_i = (\sum_{j=1}^{m_i} y_{ij}) / m_i = \hat{y}_i / M_i$ 簡單隨機取樣使

用第一階段取樣，族群總數的無偏估值推估 $\hat{\tau} = \frac{N}{n} \sum_{i=1}^n \hat{y}_i$ (公式2) σ_i^2 的變異數

$$\text{var}(\hat{\tau}) = N(N-n) \frac{\sigma_\mu^2}{n} + \frac{N}{n} \sum_{i=1}^N M_i (M_i - m_i) \frac{\sigma_i^2}{m_i} \quad (\text{公式3})$$

$$(\text{公式4}) \quad i=1, \dots, N, \quad \sigma_i^2 = \left(\frac{1}{M_i - 1} \right) \sum_{j=1}^{M_i} (y_{ij} - \mu_i)^2 \quad (\text{公式5})。$$

σ_i^2 變異數的無偏估值推估藉由取樣變異數置換族群變異數而獲得

$$\text{var}(\hat{\tau}) = N(N-n) \frac{s_\mu^2}{n} + \frac{N}{n} \sum_{i=1}^N M_i (M_i - m_i) \frac{s_i^2}{m_i} \quad (\text{公式6}) \quad s_\mu^2 = \frac{1}{n-1} \sum_{i=1}^n (\hat{y}_i - \hat{\mu}_1)^2 \quad (\text{公式7})$$

$$i=1, \dots, N \quad s_i^2 = \left(\frac{1}{m_i - 1} \right) \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2 \quad (\text{公式8}) \quad \hat{\mu}_1 = (1/n) \sum_{i=1}^n \hat{y}_i。$$

族群平均的推估， $\hat{\mu}_1 = \hat{\tau} / N$ 是每一第一單元 \bar{y}_i 族群平均的無偏估值推估，

上述變異數表示以 來區劃， $\hat{\mu} = \hat{\tau}/M$ 是每一第二單元 族群平均的無偏估值推估，變異數表示以 來區劃。

例13.1 簡單逢機取樣由族群 個第一單元中選取 個第一單元。從每一被選取的第一單元，逢機選取 個第二單元。三個被選取的第一單元的大小是 ， 和 個第二單元。樣本中第一個第一單元的y值是8和12，樣本中第二個第一單元的y值是0和0，樣本中第三個第一單元的y值是1和13，每一被選取的第一單元的總數推估 $\hat{y}_1 = \frac{24}{2}(8+12) = 240$

$$\hat{y}_2 = \frac{20}{2}(0+0) = 0 \quad \hat{y}_3 = \frac{15}{2}(1+3) = 30 \quad \text{族群總數的推估}$$

$$\hat{\tau} = \frac{100}{3}(240+0+30) = 100(90) = 9000 \quad \text{每一第一單元平均的推估} \quad \text{第一}$$

單元總數間的樣本變異數

第一單元內的樣本平均 ，，。第一單元內的

樣本變異數 ，，。的變異數推估

標準差推估7441。

比例推估

族群總數的比例推估是基於第一單元的大小 $\hat{\tau}_r = \hat{\tau}M$ $\hat{r} = \frac{\sum_{i=1}^n \hat{y}_i}{\sum_{i=1}^n M_i}$

這個推估的均方差或變異數公式的近似式

$$\text{var}(\hat{\tau}_r) \approx \frac{N(N-n)}{n(N-1)} \sum_{i=1}^N (y_i - M_i \mu)^2 + \frac{N}{n} \sum_{i=1}^N M_i (M_i - m_i) \frac{\sigma_i^2}{m_i}。$$

$$\hat{\tau} \text{ 的變異數推估 } \text{var}(\hat{\tau}_r) = \frac{N(N-n)}{n(N-1)} \sum_{i=1}^N (\hat{y}_i - M_i \hat{r})^2 + \frac{N}{n} \sum_{i=1}^N M_i (M_i - m_i) \frac{s_i^2}{m_i}$$

例13.2 例13.1的族群總數的比例推估，在族群第二單元

$$\hat{\tau}_r = \frac{240+0+30}{24+20+15}(2500) = 11,441$$

2. 第一單元的選取是取樣機率隨著樣本大小成比例

另外考慮二階段取樣設計，第一單元藉由取樣機率隨著樣本大小成比例可置換取樣來選取，第二單元的取樣是每次從第一單元中以獨立簡單逢機取樣無置換

選取。因此第二階段取樣是獨立的，任一第二單元出現在樣本中更多一次，因為第一階段是可置換取樣。

這種取樣設計，族群總數的無偏估值推估 $\hat{\tau}_p = \frac{M}{n} \sum_{i=1}^n \frac{\hat{y}_i}{M_i} = \frac{M}{n} \sum_{i=1}^n \bar{y}_i$

$\bar{y}_i = (1/m_i) \sum_{j=1}^{m_i} y_{ij}$ 取樣第 i 個第一單元內樣本平均 $\hat{y}_i = M_i \bar{y}_i$ 變異度

$\text{var}(\hat{\tau}_p) = \frac{M}{n} \sum_{i=1}^N M_i (\mu_i - \mu)^2 + \frac{M}{n} \sum_{i=1}^N \left[\frac{(M_i - m_i)}{m_i(M_i - 1)} \sum_{j=1}^{M_i} (y_{ij} - \mu_i)^2 \right]$ 變異數無偏估值推

估 $\text{var}(\hat{\tau}_p) = \frac{M^2}{n(n-1)} \sum_{i=1}^n (\bar{y}_i - \hat{\mu}_p)^2$ $\hat{\mu}_p = \hat{\tau}_p / M$ 。

2. 置換的多層級取樣設計

可置換的取樣機率隨著樣本大小成比例取樣設計對變異數的推估是很特別的。實際上，任一由已知描繪選取機率 P_i 的可置換選取第一單元的多層級取樣設計的變異數推估是相等地，對於任一被選取第一單元 i 的總和，次樣本的獨立選取在不同第一單元和無偏估值推估 \hat{y}_i 是可利用的。族群總數 τ 的無偏估值推估

$\hat{\tau}_p = \frac{1}{n} \sum_{i=1}^n \frac{\hat{y}_i}{p_i}$ 前述取樣機率隨著樣本大小成比例， $\hat{y} = M_i \bar{y}_i$ ， $p_i = M_i / M$ 。

這推估變異數的無偏估值推估 $\text{var}(\hat{\tau}_p) = \frac{1}{n(n-1)} \sum_{i=1}^n \left(\frac{\hat{y}_i}{p_i} - \hat{\tau} \right)^2$

由於可置換第一單元的選取和不同第一單元獨立的選取，這結果來自 n 個獨立的選擇。因此 \hat{y}_i / p_i 是來自某些分布的逢機取樣，用樣本的變異數來推估分布的變異數。無論多少階層的取樣設計這取樣結果被維持著。

雖然藉由這樣的變異數推估是簡單的，實際變異數推估是依賴每一階段的特定設計及每一階段取樣的條件，可藉由前述取樣設計的例證。

4. 取樣成本和取樣大小

實際利用二階段取樣，相對於相同數量的第二單元的簡單逢機取樣，觀測群團中很多第二單元比起逢機觀測遍及整個族群相同數量的第二單元常常是較容易且較不昂貴。考慮簡單逢機取樣 n 個第一單元的無偏值推估及從第 i 個第一單元簡單逢機取樣 m_i 個第二單元，單純地考慮第一單元全是相同大小， $M_i = \bar{M}$ 對於全部 i 及每一被選取第一單元的次樣本大小是 m 個第二單元。

假定取樣的平均成本藉由成本函數來描述 $C = c_0 + c_1 n + c_2 nm$ C 是調查總經費， c_0 是固定經常費， c_1 選取每一第一單元的費用， c_2 選取每一第二單元的費用。取樣中第二單元的數量是 nm 。固定成本 C ，以 $\text{var}(\hat{\tau})$ 的最小值來計算次樣本

大小 $m_{opt} = \sqrt{\frac{c_1 \sigma_w^2}{c_2 \left(\sigma_b^2 - \frac{\sigma_w^2}{M} \right)}}$ σ_b^2 是第一單元間平均的變異數， $\sigma_b^2 = \frac{\sum_{i=1}^N (\mu_i - \mu)^2}{N-1}$

σ_w^2 是第一單元內平均的變異數， $\sigma_w^2 = \frac{1}{N} \sum_{i=1}^N \sigma_i^2$ 假如 σ_b^2 沒有大於 σ_w^2 / M ，

$m_{opt} = \bar{M}$ 。 m 的最佳選擇，能藉由解 n 成本方程式獲得， $n = (C - c_0) / (c_1 + c_2 m_{opt})$

m_{opt} 的導出使用 Lagrange 方法求得。由相同第一單元和次樣本大小，

$\sigma_b^2 = \sigma_\mu^2 / \bar{M}^2$ ，簡單逢機取樣下 $\hat{\tau}$ 的變異數

$\text{var}(\hat{\tau}) = N(N-n)\bar{M}^2\sigma_b^2/n + N^2\bar{M}(\bar{M}-m)\sigma_w^2/(nm)$ $V(n, m) = \text{var}(\hat{\tau})$ 定義函數 $F = V - \lambda(c_0 + c_1n + c_2nm)$ ， λ 是Lagrange變數，函數F是對於n, m和 λ 做微分，使之趨近於零，來求解m和n。

求算在最小成本下 $\text{var}(\hat{\tau})$ 的特定值，在每一被選取第一單元的次樣本大小的最佳選擇是 m_{opt} 。變異數公式 $m = m_{opt}$ ，求解n。

對於不等大小第一單元的逢機取樣和第一單元取樣機率與樣本大小成比例有更多複雜的最佳次樣本大小選取的公式被使用，Hedayat and Sinha(1991)描述更多真實求解最適大小的問題，來計算自然現象的取樣大小n和m。

5. 多層級取樣設計的誘導

在每一階段的取樣設計，兩個基本理由對於二階段取樣設計是有條件的期望值和推估變異數的性質。T表示某些推估。如族群總數的推估， s_1 是取樣中第一單元的組合。依賴取樣被選取，T和 s_1 兩者是逢機的。第一個理由，是根據推估的期望值 $E[E(T | s_1)] = E(T)$ (公式9) T的條件期望值藉由第一單元的選取是由第二單元所有可能的次選擇替代那些第一單元。無條件期望值藉由考慮所有可能的第一單元而獲得。第二個理由是根據T的變異數

$\text{var}(T) = \text{var}[E(T | s_1)] + E[\text{var}(T | s_1)]$ (公式10) 第一項目是右邊部分包含第一單元間的變異數，第二項目包含第一單元內的變異數，可能是條件平均，藉由考慮被選取第一單元所有可能的第二單元的次樣本而獲得。無條件變異數和期望值被所有可能的第一單元的樣本逆所替代。

無偏估值推估

考慮二階段取樣設計，每一階段都是使永簡單逢機取樣和 $\hat{\tau}$ 的無偏估值推估。第一單元的樣本 s_1 ，在實際取樣總數 y_i 第i個第一單元的條件期望值推估是 \hat{y}_i ，因為在第二階段第一單元i內的簡單逢機取樣。 $E(\hat{y}_i | s_1) = y_i$ 。 $\hat{\tau}$ 的條件期望值

$E(\hat{\tau} | s_1) = E\left(\frac{N}{n} \sum_{i=1}^n \hat{y}_i\right) = \frac{N}{n} \sum_{i=1}^n y_i$ 。無條件期望值

$E(\hat{\tau}) = E[E(\hat{\tau} | s_1)] = E\left(\frac{N}{n} \sum_{i=1}^n y_i\right) = \tau$ 因為第一階段是簡單逢機取樣。

$\hat{\tau}$ 的條件期望值的變異數 $\text{var}[E(\hat{\tau} | s_1)] = \text{var}\left(\frac{N}{n} \sum_{i=1}^n y_i\right) = N(N-n)\frac{\sigma_\mu^2}{n}$ 因為第一階段的第一單元是簡單逢機取樣。

\hat{y}_i 的條件變異數，藉由第一階段選取 s_1 $\text{var}(\hat{y}_i | s_1) = M_i(M_i - m_i)\frac{\sigma_\mu^2}{m_i}$ 因為 \hat{y}_i 是第一單元i內在簡單逢機取樣下總數 y_i 的推估。因此，條件變異數 $\hat{\tau}$ ，因為由不同第一單元的第二階段選取是獨立的

$\text{var}(\hat{\tau} | s_1) = \text{var}\left(\frac{N}{n} \sum_{i=1}^n \hat{y}_i\right) = \left(\frac{N}{n}\right)^2 \sum_{i=1}^n \text{var}(\hat{y}_i | s_1) = \frac{N^2}{n^2} \sum_{i=1}^n M_i(M_i - m_i)\frac{\sigma_i^2}{m_i}$

第一單元全部可能的樣本去獲取期望值的變異數，使取樣第i個第一單元等於指標變數 z_i 且等於零，否則條件變異數 $\text{var}(\hat{\tau} | s_1) = \frac{N^2}{n^2} \sum_{i=1}^N M_i(M_i - m_i)z_i\frac{\sigma_i^2}{m_i}$

因此簡單逢機取樣下第一單元的機率 $E(z_i) = n/N$ ，條件期望值變異數

$$E[\text{var}(\hat{\tau} | s_1)] = \frac{N}{n} \sum_{i=1}^n M_i (M_i - m_i) \frac{\sigma_i^2}{m_i} \quad \text{結合二個條件根據公式10}$$

$$\text{var}(\hat{\tau}) = N(N-n) \frac{\sigma_\mu^2}{n} + \frac{N}{n} \sum_{i=1}^n M_i (M_i - m_i) \frac{\sigma_i^2}{m_i}$$

比例推估

因為二階段取樣設計的每一階段藉由簡單逢機取樣來做比例推估， \hat{y}_i 的條件期望值是 y_i ，因為第二階段為簡單逢機取樣。因此，比例推估的近似變異數是變異數的第一組成，當第二組成是基於第二階段是簡單逢機取樣維持不變時。

取樣機率與樣本大小成比例取樣

當第一單元是可置換取樣機率與樣本大小成比例取樣時， $\hat{\tau}_p$ 的條件期望值

$$E(\hat{\tau}_p | s_1) = \frac{M}{n} \sum_{i=1}^n \frac{1}{M_i} E(y_i | s_1) = \frac{M}{n} \sum_{i=1}^n \frac{Y_i}{M_i} = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{p_i} \quad p_i = M_i/M \text{ 是第一單元 } i \text{ 的}$$

選取機率。由於條件期望值來自 τ 無偏估值和變異數的第一條件的

Hansen-Hurwitz推估。變異數的第二組成，條件變異數

$$\text{var}(\hat{\tau}_p) = \text{var}\left(m_n \sum_{i=1}^n \bar{y}_i | s_1\right) = \frac{M^2}{N^2} \sum_{i=1}^n \left(\frac{M_i - m_i}{M_i}\right) \frac{\sigma_i^2}{m_i} \quad \text{因為取樣中第一單元 } i, \bar{y}_i \text{ 是簡}$$

單逢機取樣 m_i 的樣本平均。透過第一單元取樣機率與樣本大小成比例的取樣，

去獲得條件變異數的期望值。 u_i 表示第一單元 i 被選取的次數。逢機變數 u_i 是期

望值 $E(u_i) = np_i = nM_i/M$ 的二次式 $\text{var}(\hat{\tau}_p) = \frac{M^2}{n^2} \sum_{i=1}^n \left(\frac{M_i - m_i}{M_i}\right) \frac{u_i \sigma_i^2}{m_i}$ 。期望值是

$$E[\text{var}(\hat{\tau}_p)] = \frac{M}{n} \sum_{i=1}^N \left[\frac{M_i - m_i}{m_i(M_i - 1)} \sum_{j=1}^{M_i} (y_{ij} - \mu_i)^2 \right] \quad \text{變異數第二階段的組成。}$$

多階段取樣

多階段取樣設計，期望值和變異數推估能更進一步分為細目。第三階段的取樣設計，公式10的條件變異數可分解成

$$\text{var}(\hat{\tau} | s_2) = \text{var}[E(\hat{\tau} | s_1, s_2) | s_1] + E[\text{var}(\hat{\tau} | s_1, s_2) | s_1] \quad s_2 \text{ 是樣本的第二單元。} E_1 \text{ 表示}$$

s_1 的條件期望值， E_{12} 表示 s_1 和 s_2 的條件期望值，相同的條件變異數可分解為

$$\text{var}(\hat{\tau}) = \text{var}[E_1(\hat{\tau})] + E\{\text{var}[E_{12}(\hat{\tau})]\} + E\{E_1[\text{var}_{12}(\hat{\tau})]\} \quad \text{取樣設計每增加一階段，最後}$$

的條件更進一步的分解，以增加變異數的更多組成。

第十四章 雙重取樣(Double Sampling)

在比例和回歸推估的較早討論，對全體族群輔助變數的值假定是已知的。在傳統分層取樣中，假定已知族群每一單元的輔助訊息必須分類成層。有些情況，輔助變數的值，像那些有興趣的變數，只能經由取樣獲得。假如輔助值比有興趣變數值更容易或更便宜獲得，那觀測較大樣本的輔助變數比觀測有興趣的變數值更為方便。

雙重取樣的設計是首先選取單元的樣本以獲得輔助訊息，然後第二次取樣以觀測除了輔助訊息之外有興趣變數。第二次取樣選取常常選取第一次取樣的次樣本。雙重取樣的目的是藉由利用輔助變數和有興趣變數之間的相互關係去獲得較好的推估。

調查以推估林分內的樹木材積，"眼球"藉由訓練有素的觀測者可以很適當的由林分內樹木的大面積取樣去推估材積，正確的材積量測需要對這些樹木小的次樣本有所限制。在一些動物或植物的豐富度調查，可以藉由從空中粗略計算大面積的地被狀況獲得正確的計算。有些例子，正確測量有興趣變數，藉由粗略但容易去獲得輔助資料的量側。分層雙重取樣，起始樣本的單元被分類成層。第二次取樣從起始樣本分層取樣中被選取。這是很重要特殊使用的方法來調整調查中非應變數的部分。

圖14.1 雙重取樣， x_i 被觀測，由簡單隨機取樣觀察 $n' = 60$ 個單元， y_i 被增加觀測，由次樣本觀察 $n = 20$ 個單元。

1. 以雙重取樣做比例推估

第 i 單元有興趣變數經常一用 y_i 來表示，輔助變數用 x_i 來表示。第一次取樣的單元數量以 n' 來表示，第二次取樣的單元數量以 n 來表示。第二次取樣的每一單元， x_i 和 y_i 兩者都被觀察，第一次取樣的其餘單元，只有 x_i 被觀察。

假定 y_i 和 x_i 是高度相關，藉由線性相關，當 $x_i=0$ ， $y_i=0$ 。以雙重取樣做比例推估可以改進族群平均或總和的推估。

假定從族群 N 個單元， n' 個單元樣本是由無置換簡單隨機取樣被選取觀測的 x 值。從被選取的 n' 個單元，亦由無置換簡單隨機取樣選取 n 個次樣本被選擇做 y 值的觀察，以增加對 x 值的觀測。

從包含 x 和 y 值兩者的小樣本，獲得樣本的比例 $r = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n x_i}$ 總數包括第二次

樣本的所有單元。從觀察 x 值的全部樣本，能推估 x 值的族群總數 τ_x $\tau_x = \frac{N}{n'} \sum_{i=1}^{n'} x_i$

總數包括起始樣本的所有單元。 y 值的總數 τ 的比例推估 $\hat{\tau}_r = r \hat{\tau}_x$ 。這種均方差

或變異數推估的近似方程式 $\text{var}(\hat{\tau}_r) \approx N(N-n') \frac{\sigma^2}{n'} + N^2 \left(\frac{n'-n}{n'} \right) \frac{\sigma_r^2}{n}$ σ^2 是

有用有限族群 y 值的變異數， σ_r^2 是族群比例變異數 $\sigma_r^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - R x_i)^2$ 族群比例

$R = \tau/\tau_x$ 。變異數推估 $\text{var}(\hat{\tau}_r) = N(N-n')\frac{s^2}{n'} + N^2\left[\frac{n'-n}{n'n(n-1)}\right]\sum_{i=1}^n(y_i - rx_i)^2$ s^2 是

次樣本y值的樣本變異數。

例14.1 麋鹿豐富度調查，總數240隻麋鹿在航測調查的20個樣區被發現。當那些樣區中的5個做地面調查—56隻麋鹿在航測調查中被發現—總數70隻麋鹿被發現。研究區域包括100個大小相同的樣區。推估研究區域的麋鹿數量。

有興趣變數y是樣區中實際麋鹿的數量，藉由地被狀況來推測。從空中進行觀測—較粗略但能涵蓋樣區中的較大數量—輔助變數x。從來自空中和地面觀測的次樣本 $n = 5$ ，樣本的比例 $r = \frac{70}{56} = 1.25$ ，從空中觀測全部 $n' = 20$ 的樣本，

族群總數的推估 $\hat{\tau}_x = \frac{100}{20}(240) = 1200$ ，研究區域麋鹿數量的比例推估，基於雙重取樣設計 $\hat{\tau}_r = r\hat{\tau}_x = 1.25(1200) = 1500$ ，因此研究區域麋鹿總數量的推估是1500。

這類型的取樣規劃，樣本比例的倒數是偵測的推估。因此， $56/70 = 1/1.25 = 0.8$ 是從空中偵測被選取樣區的任一麋鹿的偵測推估。

2. 雙重取樣比例推估的配置

在有些狀況比例推估的雙重取樣是有關連的，有興趣y的變數對於輔助變數x傾向於線性相關，當 $x=0$ ，y趨向於0，且測量x比y更容易或更便宜。在雙重取樣中對於次樣本的理想比例的取樣依賴著觀測兩個變數的相關成本和雙重取樣比例相關的強度。假定觀察一個單元x值的費用是 c' ，觀察y值的費用是 c ，所以總成本C是 $C = c'n' + cn$ 固定成本C的最小變異數推估 $\hat{\tau}_r$ 是由次樣本取樣中獲得

$$\frac{n}{n'} = \sqrt{\frac{c'}{c} \left(\frac{\sigma_r^2}{\sigma^2 - \sigma_r^2} \right)}。$$

3. 分層雙重取樣

有些取樣的情況，在樣本被選取之後單元被歸於分層中。例如，單單電話抽樣後，藉由逢機電話的人數選取樣本將性別，年齡或職業分層。在逢機取樣調查魚類去推估平均年齡也許可以一選取的大小等級分層。事後分層的方法在第11章分層取樣中被描述，使用的狀況是對於每一層h層的族群單元相對比例 $W_h = N_h/N$ 是已知的。假如族群適當比例不知，雙重取樣也許可用，藉由起始(大)樣本將單元分層和從起始樣本做分層取樣選取。雙重取樣是可以利用的，假如層內的單元是相類似的及假定輔助性態值分層是基於測量較測量有興趣變數更容易或較不昂貴。

從族群N單元中起始簡單逢機取樣選取 n' 個單元。這些單元被分類成層，h層觀測 n'_h 個單元， $h=1, \dots, L$ 。h層單元的族群比例 W_h 是藉由樣本比例 $w_h = n'_h/n'$ 來推估。第二次取樣是藉由從第一次取樣分層逢機取樣來選取， n_h 單元的選取是從h層 n'_h 樣本單元中選取，由第二次取樣的每一單元有興趣變數 y_i 被記錄。在第二次取樣h層的樣本平均是 $\bar{y}_h = \sum_{i=1}^{n_h} y_i / n_h$ 。族群平均的推估 $\bar{y}_d = \sum_{h=1}^L w_h \bar{y}_h$ 。 \bar{y}_d 推

估是族群平均 μ 的無偏估值，變異數 $\text{var}(\bar{y}_d) = \frac{N-n'}{Nn'}\sigma^2 + \sum_{h=1}^L \frac{W_h\sigma_h^2}{n'}\left(\frac{n'_h}{n_h}-1\right)$

(公式1) σ_h^2 是h層內族群變異數， σ^2 是整個族群的變異數。公式1的第一項

是起始取樣 \bar{y}_d 的期望值變異數。假如全部起始取樣的y值被記錄及樣本平均被推估，變異數將可獲得。第二項是期望值，全部可能的起始樣本，起始樣本分層取樣平均 \bar{y}_d 的條件變異數。第二階段取樣部分 n_h/n'_h 假定是固定的且樣本 $n'_h=0$ 是排除的。雙重取樣 \bar{y}_d 的變異數無偏估值推估，利用從第二取樣分層取樣變異數 s_h^2 是 $\text{var}(\bar{y}_d) = \left(\frac{N-1}{N} \right) \sum_{h=1}^L \left(\frac{n'_h-1}{n'_h-1} - \frac{n_h-1}{N-1} \right) \frac{w_h s_h^2}{n_h} + \left[\frac{N-n'}{N(n'-1)} \right] \sum_{i=1}^L w_h (\bar{y}_h - \bar{y}_d)^2$ (公式 2)。

4. 調整調查中非應變數

雙重取樣對於分層的重要使用是解決調查中非應變數的調整回收。在人類族群的調查，非應變數——人被取樣選取不是沒有利用或者不願回答——常常不是代表有興趣變數。在一些情況，當族群平均是推估時，樣本平均的應變數是具偏值的。相同地，自然族群的調查，藉由一個理由或另一個不被觀察因子來選擇地點對於整個研究區域是不具代表性的。例如，魚類族群的拖網調查，當海洋地形有太多岩石或其他干擾拖網的因子時，有些位置沒有被觀察到。在有些地點棲息地和魚類種類組成也許不是典型的族群，因此拖網地點的樣本平均對於族群平均的推估將是具偏估值推估。

有些例子非應變數單元也許考慮分離——但單元是取樣選取之後未知如何分層。非應變數層的次樣本被利用，經常需要加強其效果或在起始取樣採用不同的方法。在電話訪問的例子，次樣本包括回收或訪問員拜訪。在魚類調查的例子，對於難到的地點採用不同的魚網或其他的取樣方法或許是必須的。

讓分層1包括應變數的單元，分層2包括非應變數的單元。起始簡單取樣選取 n' 單元。在調查的起始部分， n'_1 是應變數單元的數目和 n'_2 是非應變數單元的數目。在回收或二次效果，應變數是藉由起始非應變數 n_2 的簡單逢機取樣而獲得。相關連地，第二階段的分層取樣，包括全部起始應變數 n'_1 及起始 n'_2 非應變數 n_2 。

族群平均的無偏估值推估藉由分層雙重取樣來推估 $\bar{y}_d = \frac{n'_1}{n'} \bar{y}_1 + \frac{\bar{y}_2}{n'} \bar{y}_2$ ，變異數，由公式1， $\text{var}(\bar{y}_d) = \frac{N-n'}{Nn'} \sigma^2 + \frac{n'_2-n_2}{n'n_2} W_2 \sigma_2^2$

5. 雙重取樣的統計值

近似平均和變異數，比例推估

當二階段取樣設計，雙重取樣的推估是評估第一次取樣的條件。 s_1 是表示起始選取 n' 個單元，使 s_1 ， $\hat{\tau}_r$ 推估是基於 s_1 的 n' 個單元簡單逢機取樣 n 個樣本的比例推估，所以條件期望值近似 $E(\hat{\tau}_r | s_1) \approx N \bar{y}_{s_1}$ ， \bar{y}_{s_1} 是 s_1 的 n' 個單元的y值平均。因此 \bar{y}_{s_1} 是簡單逢機取樣的樣本平均，無條件期望值的近似 $E(\hat{\tau}_r) \approx E(N \bar{y}_{s_1}) = \tau$ 及

$\text{var}[E(\hat{\tau}_r | s_1)] \approx \text{var}(N \bar{y}_{s_1}) = N^2 \left(\frac{N-n'}{Nn'} \right) \sigma^2$ ，由 s_1 做比例推估 $\hat{\tau}_r$ 的條件變異數的近似式 $\text{var}(\hat{\tau}_r | s_1) \approx \frac{N^2(n'-n)}{n'n(n'-1)} \sum_{i=1}^{n'} (y_i - r_1 x_i)^2$ ， r_1 是藉由 s_1 的 n' 個單元的x值總數區劃y值總數。在簡單逢機取樣下比例推估的變異數推估，所有可能選取的 n' 個單元近似式 $E[\text{var}(\hat{\tau}_r | s_1)] \approx \frac{N^2(n'-n)}{n'n} \sigma_r^2$ $\hat{\tau}_r$ 的變異數推估

$$\text{var}(\hat{\tau}_r) = \text{var}[E(\hat{\tau}_r | s_1)] + E[\text{var}(\hat{\tau}_r | s)] \quad \text{var}(\hat{\tau}_r) \approx N(N - n') \frac{\sigma^2}{n'} + N^2 \left(\frac{n' - n}{n'} \right) \frac{\sigma_r^2}{n}$$

這種推估使用無偏估值 s^2 來推估 σ^2 及比例變異數推估來推估 σ_r^2 。

比例推估的最佳配置

藉由 $\text{var}(\hat{\tau}_r)$ 最小值找尋 n' 和 n 是受限於 $c'n' + cn = C$ 。使用Lagrange方法求解。 $f(n', n) = \text{var}(\hat{\tau}_r)$ $f(n', n) = N^2 [n'^{-1}(\sigma^2 - \sigma_r^2) + n^{-1}\sigma_r^2 - N^{-1}\sigma^2]$

$g(n', n) = c'n'cn - C$ 定義 $H(n', n, \lambda) = f - \lambda g$ 求解

$$\frac{\partial H}{\partial n'} = \frac{-N^2(\sigma^2 - \sigma_r^2)}{n'^2} - \lambda c' = 0 \quad \frac{\partial H}{\partial n} = \frac{-N^2(\sigma_r^2)}{n^2} - \lambda c = 0 \quad \text{忽略Lagrange變數}$$

λ 藉由分離第一方程式 $N^2 c'$ 及第二方程式 $-N^2 c$ 和增加二個方程式

$$\frac{\sigma^2 - \sigma_r^2}{c'n'^2} = \frac{\sigma_r^2}{cn^2} \quad \frac{n}{n'} = \sqrt{\frac{c'}{c} \left(\frac{\sigma_r^2}{\sigma^2 - \sigma_r^2} \right)}$$

J. N. K. Rao(1973)提及分層雙重取樣的結果及小部分最佳取樣的結果，Cochran(1977)和Singh和Chaudhary(1986)進一步討論分層雙重取樣，

第二十六章 分層適應群團取樣

(Stratified Adaptive Cluster Sampling)

在分層適應群團取樣中，起始分層取樣由族群中被選取，且任一單元有興趣變數值被觀察以滿足特定的條件，從鄰近地區增加取樣單元到樣本中。假如隨後增加的單元能符合條件仍有更多的單元增加至樣本中。

從實際觀點而言，分層適應群團取樣設計是重要的，因為很多族群重要資料存在於起始分層中且族群集中的形式或實際分布並不能預測。在傳統分層取樣中，認為相似的單元被集合分類成層，基於族群重要的資訊或相似的單元。適應群團取樣，另一方面提供族群利用群團趨勢的方法，在群團位置和形狀在調查之前不能預測時。在這一章節取樣設計結合二種方法。

傳統推估像分層取樣平均值用適應性設計是具偏估值的，在這章節將提及適應性取樣設計下的推估具偏估值。分層適應群團取樣複雜化提升是從分層選可能取從其他分層增加單元進入樣本，所以分層觀察不是像傳統分層取樣是獨立的。在本章不同無偏估值推估使用跨越層邊界的不同細微地方法來求算。本章很多結果引自Thompson(1991b)。

1. 取樣設計

本章分層適應群團取樣，族群被劃分成L層，h層是由 N_h 單元所組成，族群單元總數是N。關於 u_{hi} 單元，h層的i單元， y_{hi} 是有興趣的變數。族群任一 u_{hi} 單元， u_{hi} 單元的鄰近單元被定義為包括 u_{hi} 單元的取樣收集單元且是有意義的特性，假定 $u_{h'i'}$ 單元是 u_{hi} 單元的鄰近單元，那 $u_{h'i'}$ 單元的鄰近單元是 u_{hi} 單元。鄰近單元可能包括一個分層更多的單元。 u_{hi} 單元是滿足有興趣變數的條件，假如y值在特定組合C中與單元相關連，

在這章取樣設計的考量，使用分層逢機取樣從族群起始取樣選取單元；h層內，簡單逢機取樣無置換選取 n_h 單元，從分離層中是獨立選取。每當被選取單元符合條件，鄰近不是在樣本中的所有單元被增加至樣本中。當被增加的任一單元符合條件時，仍有很多單元被增加至樣本中，所以最後樣本包括任一取樣單元符合條件的鄰近地區的每一單元。

實例在圖26.1中說明，目的是要推估群團族群的豐富度，總數經由每一單元點目標的y數量的區域單元來求算。點目標舉例來說，代表植物或動物的位置。單元符合條件假定包括一個或多個點目標， $y \geq 1$ 。在圖26.1顯示，族群被劃分成兩層並由每一層以簡單逢機取樣選取5個單元。單元的鄰近地區包括單元全部北、南、東和西的鄰近單元。應用分層適應群團取樣的最終樣本以圖26.2表示。

族群也許被劃分成K組單元組合，稱為網狀，在網狀中起始取樣的任一樣本被選擇包括網狀最終樣本的所有單元。單元不符合條件屬於網狀本身的組合。符合條件的起始選擇單元表示增加單元到樣本結果，不只網狀所有其他單元還包括網狀的鄰近區域或更多數量的網狀。在圖26.3，起始樣本交叉網狀以粗線條表示。在網狀交叉起始取樣中，樣本的其他單元“邊緣單元”沒有符合條件且不在起始樣本中，但是每一鄰近區域的一個或多個單元符合條件。

$$E(r_{hi}) = \sum_{k=1}^L n_k \frac{m_{khi} + a_{khi}}{N_k}$$
$$\pi_{hi} = 1 - \prod_{k=1}^L \binom{N_k - m_{khi} - a_{khi}}{n_k} \bigg/ \binom{N_k}{n_k}$$

2. 推估

3. 與傳統分層取樣做比較

4. 進一步改進推估

5. 例子

地理資訊系統於分層取樣調查設計與分析上之應用
Geographic Information System Used to Stratified Sampling
Design and Analysis

1. 地理資訊系統之空間圖層資料
2. 地理資訊系統之空間分析套疊分析
3. 分層取樣調查設計
 事前分層
4. 取樣調查資料分析與展示
 事後分層

(一) 分層取樣調查設計

以往的資源調查雖名為雙重取樣，但其實只是利用簡單逢機 (sample random) 或系統取樣 (systematic field sampling) 以求得航空照片樣點與地面樣區，再分別由航空照片判釋材積和地面調查求取樣區材積以為蓄積推測之依據而已，並沒有將航測資料與地面樣區資料加以整合，以改進材積推估能力。

另外有關分層問題，假如分層 (stratification) 成功則樣區或樣點可以充分代表不同族群，若以相同的經費來加以比較，則分層取樣將可獲得較精確的材積。然分層取樣可分為事前分層與事後分層，亦即調查前、後依不同林分性質進行分層，以利調查或分析，台灣三次資源調查皆係採事後分層。

分層可由圖或像片加以取樣，然一般由航空照片來得多，整合照片樣區與地面樣區為分層取樣設計，為較省錢的工作。然據研究由蓄積級分層是無法減少推算生長收穫的變異 (variance)，即依蓄積級分層並無法改善蓄積推算值 (Maclean, 1972)；其精確性 (precision) 則完全依照取樣數量來決定。

1、分層選點

森林分層可依 (1) 地形 (topographical feature)

(2) 地位級 (site classes)

(3) 坡度 (aspect)

(4) 坡向 (slope)

(5) 林型 (forest type)

(6) 密度級 (density classes)

(7) 材積 (volume)

(8) 樹高

(9) 年齡

(10) 海拔高 (altitude)

(11) 事業區 (working circle)

(12) 集水區 (watershed)

分層之基礎盡可能以推估時相同性態值為之，目前改進以地理資訊系統之不同圖層(map layers) 重疊以獲所選定之區劃範圍。

2、分層取樣

分層(stratification)簡單講即為等級的使用，應用於取樣即為在同一等級裡其個體皆相類似(級內的變異很小)，在其他等級裡則不同。若應用分層於森林資源調查上，則可以以較少的勞力、經費來調查樣本，而達到相同的正確度(accuracy)。

適當的使用分層將可以提高取樣的效率(efficient sampling)，何謂適當，可以分成三項來加以說明：

- (1) 各定義層(defining strata)一定要獨立取樣以獲各層級的平均值。各層要獨立取樣，意即各層級的分布必須於個體被選擇計算各層級平均值前即已知。在森林資源調查上有許多原則可供分層級，如地理因子(geographical factor)、林型(forest type)或林分狀況(stand condition)等。這些因子可由以往的圖籍、航空照片判釋或事先的地面調查等事先查知。目前最有效的過程為利用高品質的航空照片、高技術的像片判釋和遙感探測技術等以便獲得地理資訊系統內所需要的高品質資料來加以研判。
- (2) 各層級的平均值必須確實不同。此平均值係指樣本的平均值(arithmetic average or mean)與離均(平方)差(average squared deviation from mean or variance)、變異(variance)或均數的變異數或標準差(標準誤差)，各層級間的差異愈大，則分層取樣之效率愈較簡單隨機取樣來得高。
- (3) 各層級的觀測值分布一定要適當。決定適當與否有二法，一為比例分派(proportional allocation)，二為最適分派(optional allocation)。
 - A. 各層級內的變異(variance)皆相同，若這樣則依各層級的權重的比例予以分派，於資源調查中，各層級的面積是分派的權重依據。
 - B. 若層級的變異數不相同，比例分派仍然可用，但最適分派(optimum or Neyman allocation)則較有效率(即以較少的成本獲得想達成的正確度)。最適分派意為觀測值與(各級面積×標準偏差)乘積呈比例分佈(distributed proportional to the product of class area & standard deviation)。(註：標準偏差standard deviation = $\sqrt{\text{variance}}$)
 - C. 若觀測值其各層級的變異既不相同且不呈比例，則分層取樣對此族

群沒有效率。

由以上所提，一般地理因子(geographic commits)如鄉鎮、集水區可能無法由分層取樣獲益，林型會好些。

如何分層級，可由許多原則決定，但最有效的原則還是與欲推估事物有密切關連者。

林木資源調查(timber cruising)的目的一般為推算材積。在這種情況最有效的分層標準為減少同層級內材積變異，或增加層級間級平均值的差異。

(二) 永久樣區之取樣規劃實務

1、永久樣區之取樣規劃

以新化林場地理資訊系統資料庫所建立之圖層(馮等 1993, 陳 1993), 配合地理資訊系統套疊(overlay)、選擇(select)之功能, 將林地內具有不同地理條件之地點加以標示, 以不同之樹種、林齡、地理座標、海拔高土壤類型、坡度、坡向做為分層取樣, 規劃永久樣區之依據。規劃步驟如下所列:

(1) 資料蒐集

蒐集造林台帳、像片基本圖、航空照片、歷年研究文獻及數位地形模型(Digital Terrain Model, DTM)。

(2) 土地利用型圖(含林型圖)及林業主題圖之彙整

依陳(1993)所建之新化林場地理資訊圖檔(林場界、林班界、林道、水系、土地利用型、等高線、坡度、坡向、立體圖及土壤圖)為基礎, 轉換成相同網格大小之網格式圖檔, 土地利用型圖依不同屬性及其林型代號(tag), 查詢出單一土地利用型(含林型)。

(3) 圖形套疊及樣區選點

以林班界圖、土壤圖、坡度圖、坡向圖, 配合已建立之資料庫圖層分別與不同的樹種(人工林)或林型(天然林)圖, 加以套疊, 產生不同地理因子與土壤類型之林型圖, 再分別對應造林台帳決定林齡。至此, 即可於圖面上將相同之林型, 不同坡度、坡向和土壤類型之各種組合區域標示出來, 並於各區域中以網格板逢機決定樣區位置。

(4) 尋找樣區、推算調查費用

套疊林道圖、水系圖林班界圖即可獲知樣區之座標以及到達樣區所需路徑、地理狀況, 以便計算樣區調查時間、費用, 更可配合GPS或航照片、基本圖, 尋找出樣區正確位置。

永久樣區坡度、坡向分層之檢定：

以Bertalanffy模式預測各永久樣區之基準林齡(14 年生)胸高直徑生長量，再以進行非均衡性變方分析，結果顯示：坡度、坡向之群間皆具有顯著差異，但坡度、坡向間之交感效應即未檢定出變異，如表一。故可知永久樣區之設立上，分層取樣之考量仍有其具體意義。

表一 不同坡度、坡向樣本變方分析表

變因	DF	SS	MS	F Value
坡度	2	25.441956	12.720978	* 6.17
坡向	3	21.403898	7.134633	* 3.46
交互效應	6	20.492197	3.415366	1.66

*表差異顯著水準% (MSE=2.061746)

*DF 為分群後之自由度

以新化林場而言，樹種組成包括了大葉桃花心木、柚木、竹、龍眼、麻六甲合歡、相思樹、樟樹、黑板樹、芒果、桉樹、苦楝等，其中以大葉桃花心木最為重要，在永久樣區之設定上實應優先考量；若欲配合林場地理資訊系統之建立，則其地理因子更突顯出在取樣上的重要性。新化林場海拔高度在公尺至152公尺間，變化不大，坡度、坡向在本研究中證實對林木造成生長上的差異，可提供今後永久樣區設置時之參考。

在實際作業上，由於坡度、坡向二種地理因子無法在分層取樣中做細度的分級，本實驗坡向以方位角表示且 $^{\circ}$ 為一等級，若依Whittaker(1960)提出之方位與水分梯度之關係，如圖一；則亦可由水分梯度來探討其對林木生長之影響。

新化林場之地勢較平坦，坡度分級以度為一級。

<hr/>															
NE		ENE		E		ESE		SE		SSE		S		SSW	
NNE		N		NNW		NW		WNW		W		WSW		SW	
<hr/>															
1	2	3	4	5	6	7	8	9	10						

圖一 根據方位之水分梯度 (Wittaker 1960)

(三) 永久樣區之取樣規劃及設立

利用新化林場已建立之森林經營資料庫(陳1993)，做為大葉桃花心木永久樣區取樣規劃之依據，林型圖、土壤圖、坡度圖、坡向圖之套疊，可得不同

坡度級、坡向級、土壤類型之分層資料，提供樣區設置時之參考。舉例而言，坡度5-15度、坡向為北向，土壤類型為壤土之大葉桃花心木林型，其可供樣區設立之範圍，可經由各相關圖層之交集分析產生，再經由林班區界、林道圖之疊合，提供樣區之位置(如圖九)，並達成以坡度、坡向、土壤類型為分層依據之目的。因新化林場之坡度區分為 6 級，坡向區分為 8 級(依據第三次資源調查)，同時包含 5 種土壤類型，疊合工作極為繁瑣，應考慮各個分級間之合併或設計輔助軟體，以節省作業時間。

4. 取樣調查資料分析與展示

事後分層

(二) 方法

- 1、資料處理—將文字檔轉入III和LOTUS 以方便整理。
將圖檔建入地理資訊系統中處理、展示。
- 2、資料分析—將地面樣區調查之單株林木資料進行林分結構與林分蓄積分析、並將結果累彙送至林分資料檔。

(1) 樹種組成

各樣區由航空照片判釋及地面植生踏勘所判別繪製之林型為植群分類並進行各樣區內樣木的樹種種類分析以得樹種組成依性質將樣區歸屬為主要樹種林型。

整理主要土地利用型內所含主要樹種林型樣區之百分比天然林之森林植群之組成、林分結構、林分蓄積是受生育環境影響。

樹種組成—用相對頻度、相對密度、相對優勢度之和表示之重要值。

重要值指數(Importance Value Index, IVI)

- A. 為學派改善指數 (Density-Frequency-Dominance Index) 密度、頻度、優勢度(或覆蓋度)絕對值，指數大小沒有一定範圍，且三介量單位不一總和以代表的不合理現象。
- B. IVI為相對密度+相對頻度+相對優勢度，每個介量最大值為，植物重要值之總合恆為，若一植物社會值由一植物組成，則其重要值為300。
- C. 重要值可顯示某植物於該林型或林分中相對重要性，且兼顧密度、頻度及優勢度之綜合特性(劉、蘇1989)。

$$\text{相對頻度}\% = \frac{\text{某植物之頻度} \times 100}{\text{林分中所有植物頻度之總和}} \dots\dots\dots (1)$$

相對密度(relative density)%

$$= \frac{\text{某一植物之株數}}{\text{所有樣區內全部植物之株數}} \times 100 \dots\dots\dots (2)$$

$$\text{相對優勢度 (relative dominance) \%} = \frac{\text{某一植物之優勢度 (或覆蓋度)} \times 100}{\text{所有樹種優勢度之總和 (或覆蓋度)}} \dots\dots\dots (3)$$

$$\text{重要值指數 (Importance Value Index, IVI)} = \text{相對頻度} + \text{相對密度} + \text{相對優勢度} \dots\dots\dots (4)$$

(2) 林分結構

林分結構是指林地、樹種、林木各種性態值大小分布，為樹種生長習性，環境條件和經營實施的綜合結果 (Husch, and Beer 1972)。天然林的林分結構可由林木的性態值如胸高直徑、斷面積、樹高、材積大小的分布加以說明。所以直徑分布是林分結構中的一種表示方法，乃因胸高直徑資料最易獲得，故常被使用，而直徑分布以機率密度函數 (如 (5) 式) 描述得最理想 (馮1988)。

A. 各樣區內樣木的林分結構以機率密度函數 (x, θ) 來配合直徑分布，求算出的三個母數 a, b, c 之最佳推估值 $(a, b, c,)$ 。

$$f(x, \theta) = (c/b) [(x-a)/b]^{c-1} \exp[-(x-a)/b] C \dots\dots\dots (5)$$

式中的 $a \geq 0, b > 0, c > 0$

x 為胸高直徑

\exp 為指數

(3) 建立分析推估材積之模式系統推估林分蓄積

林分蓄積可由林分內單株林木材積累積，或直徑分布法 (diameter distribution method) 由各直徑級之材積累計，或全林分生長模式 (whole stand method) 由全林分材積與樹種、林齡、地位、密度、處理間的迴歸式推估而得。由於天然林林木種類複雜，林齡不易獲得，所以不加以考慮。環境因子對植生組成、結構、蓄積之影響，有層次且次序相關性存在，以地形位置 (topographic position) 及方位 (aspect) 為最上層因子，可解釋大部分植生之地域變異，稱為環境層級觀念 (Kellman 1980)。所以以坡度、坡向、海拔高等地理因子代表地位。密度則以單位面積之林木株樹表之。

A. 整理各重要樹種之單株林木材積式，樹高曲線式。

B. 利用單株林木材積式推估累計樣區內之林木材積 (SUV)。

以樣區面積除之，以獲得單位面積之林木蓄積 (SUV/ha)

$$SUV/ha = \sum V_i / ha \quad V_i = f(D, H) \dots \dots \dots (6)$$

式中的 V_i ：株林木材積
D：胸高直徑
H：樹高

C. 利用直徑分布法求算樣區之林木材積 (DDM)。

以樣區面積除之，以獲得單位面積之林木蓄積 (DDV/ha)

$$DDM/ha = Nt \int (x) f(x, \theta) \dots \dots \dots (7)$$

式中的 Nt 為樣區內林木株數換算之單位面積之林木株數
 $g(x)$ 為胸高直徑的函數式如 $H=f(D)$ $V_i=f(D, H)$
 $f(x, \theta)$ ：的母數 θ — a, b, c

D. 將樣區內樣木所求算之 $main\ sps$ ， a ， b ， c ， Nt/ha ， SUV/ha ， DDV/ha 轉入樣區林分檔中。

E. 繪製林分結構母數 (a, b, c) 與材積母數 (SUV/ha ， $/ha$) 與坡度 (slope)、坡向 (aspect)、海拔高 (AH)、林齡 (age) 之關係圖由此判斷林分結構、材積與環境、林齡之間的關係為直線或指數、倒數。

F. 建立各SPS土地利用型與林型、林分結構、材積與地理因子，建立樹種空間性材積推估模式。

建立林分結構、林分蓄積分析模式系統架構。

各林型 $SUV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $DDV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $\theta = f(\text{坡度、坡向、海拔高})$

G. 將土地利用型歸併為種天然林林型及包含人工林、建地等其他土地利用型。依第三次資源調查所繪製之林型圖算出各種天然林涵蓋的面積以及顯示其所在位置範圍之天然林林型圖。

H. 利用 (c) (d) 所算出之單位面積之林木蓄積 (SUV/ha) 與 (DDV/ha)，配合 (G) 算出之各種天然林涵蓋的面積，可以以比率推算 (ratio estimation) 出各種天然林型之林木蓄積。

I. 以農林航空測量所所製做之數值地形模型 (DTM)，疊合以上繪置製之各種天然林林型圖求出各種天然林涵蓋的 $40m \times 40m$ 地理單元的平均坡度、坡向、海拔高。

J. 以 (E) 所建之材積迴歸推估模式，來推算 (regression

estimation) 各種天然林型涵蓋地理單元之材積並加以累計。可得各種天然林型之林木蓄積，。

K. 以 (H) 所建之母數 (θ) 迴歸推估模式，來推算各地理單元的母數 (parameter estimation)，再藉 (d) 的直徑分布法，亦可推算天然林型涵蓋地理單元之材積並加以累計。可得各種天然林型之林木蓄積 DDVpar。

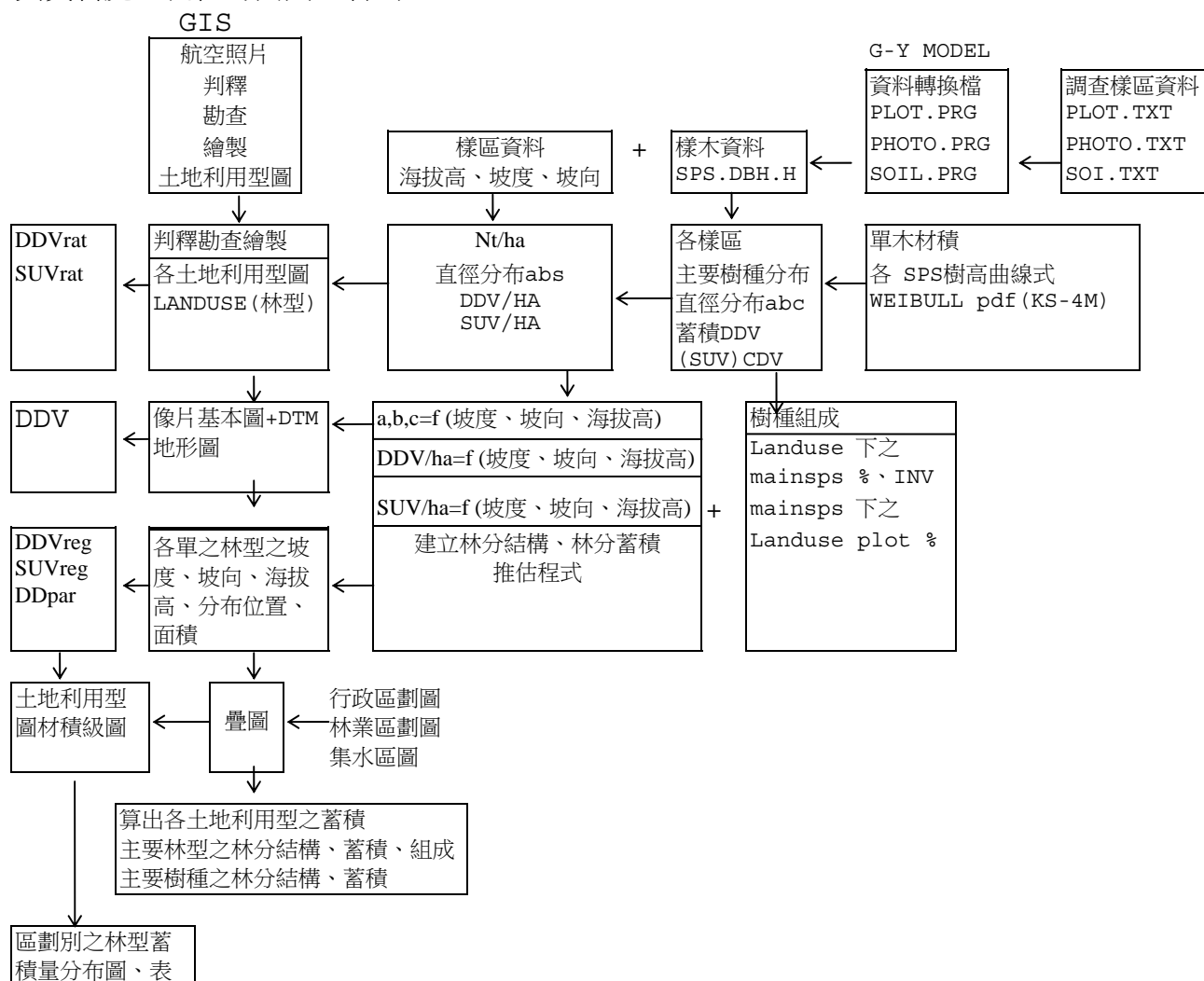
L. 由以上 (H) (J) (K) 可以算出 SSVre，DDVre，SUVreg，DDVreg，DDVpar 五種結果。

M. 評估選擇最合理之材積推估值。

N. 利用地理資訊系統圖籍資料與屬性資料關聯功能，繪製材積級圖。

O. 與林班圖、事業區圖、集水區圖等林業、行政區劃主題圖疊合，即可獲各林業、行政區劃土地利用型的各種樹種林型之林分結構及材積分布圖與數量圖。

其資料處理流程可由圖二看出：



圖二：台灣第三資森林資源調查蓄積量資料處理流程

綜合其結果可得：

(1) 林分樹種組成分析

- A. 分析調查區主要樹種、直徑級株數及於樣區中之出現率，以決定林型、樹種別。
- B. 整理各方位、海拔、經緯度下之林型、樹種別。

(2) 林分結構分析

- A. 求算各地面樣區、航照片樣區及各樹種別、林型別之胸高直徑分布與平均每公頃蓄積量。
- B. 建立各樹種別、林型別之材積推估模式。
- C. 以Weibull pdf 來配置各樣區之直徑分布，利用五種母數求解法求出Weibull 母數a、b、c。

(3) 建立分析推估材積之模式系統。

A. 單位面積蓄積量推估

(a) 合計樣區內之所有利用單木材積式計算之樣木的材積再除以樣區面積以得單位面積蓄積量之推估 (SUV/ha)。

(b) 利用直徑分布法推估單位面積之蓄積量 (DDV/ha)。

B. 建立林分材積推估模式

(a) 建立林分生長收穫隨著UTM坐標位置、海拔高、坡度、坡向等空間因子不同而改變之趨勢變化。

(b) 利用逐步迴歸建立林分材積推估模式。

C. 推算全林分材積

(a) 利用地理資訊系統 (GIS) 的疊圖分析功能將林分蓄積，結構的空間模式與數位地形 (DTM) 和土地利用型圖整合分析、推算出組成各土地利用型積結構分布圖、蓄積量分布圖。

(b) 疊合所需瞭解的行政區劃、林業區劃、集水區區劃圖層以掌握，各種區劃之樹種組成、林分結構、林分蓄積量。

(4) 繪製各種區劃之林型材積分布圖。

3、模式驗證：以第三次森林資源調查之地面樣區之林分資料檔結果驗證各主要林型之天然林林分蓄積推估模式並分析誤差來源。

七、附錄

全省尺度、林管處尺度、事業區尺度、林班尺度

1. 依土地利用型 (主要林型)、位置座標、海拔高、坡向、坡度為分層標準，將樣區分屬歸類

2. 歸屬同類樣區內之樣木，予以合併 (pooling)，計算

(1) 樹種組成：各林型主要樹種組成。

(2) 掌握林分結構 (直徑分布) — 以Wei-bull機率密度函數建立描述各樣區或各樹種之林分結構 (直徑分布) 情形，由KS適合度測驗，發現配置理想且其母數a，b，c值可解釋各直徑分布情形。更而得知各直徑級佔有株數頻度及材積分布。

(3) 蓄積量

a、關係式、材積式為單株材積林分蓄積量推估之核心。

$$H = f(D) \quad V = f(DHF)$$

b、單位面積之蓄積量可由單株材積累計單株材積、林分蓄積量 (SUV/ha)

和直徑分布法 (DDV/ha) 計算之。

c、區劃範圍之林分蓄積可分比例推估法、迴歸推估法與母數推估法三種。

3. 由GIS土地利用型圖(主要林型圖)、位置座標分區圖、海拔高圖、坡向圖、坡度圖等圖層予疊合，可算出各種分層條件下的土地利用型面積(主要林型)面積，再以樣區分析之單位面積之材積比例推算各林型之林分蓄積量。

面積及蓄積

1. 土地利用型區分

林地部份係依林地上經濟樹種之生立木中蓄積佔多數之樹種而決定。在照片上區分林型係依據各樹種樹冠所佔面之比率和樹冠大小而判斷。林分內經濟竹林佔林冠覆蓋一半以上時區分為竹林；非林地部份則依實際狀況予以區分。土地利用型用以下列三位阿拉伯數字代號表示：

代號	土地利用型	代號	土地利用型
011	冷杉天針	114	台灣杉造林
012	鐵杉天針	115	柳杉造林
013	檜木天針	116	肖楠造林
014	松類天針	119	其他針造林
015	雲杉天針	120	人針混
019	其他天針	130	人針闊混
030	天針闊混	141	相思樹造林
040	天闊純	142	楓香造林
050	天闊混	143	樟樹造林
061	桂竹林	144	光臘樹造林
062	孟宗竹林	145	台灣櫟造林
063	麻竹林	146	桐類造林
064	荊竹林	149	其他闊造林
065	綠竹林	150	人闊混
069	其他竹林	161	桂竹造林
070	天竹針混	162	孟宗竹造林
080	天竹闊混	163	麻竹造林
090	天竹針闊混	164	荊竹造林
111	檜木造林	165	綠竹造林
112	松類造林	169	其他竹林
113	杉木類造林	170	人竹針混
代號	土地利用型	代號	土地利用型
180	人竹闊混	639	其他果園
190	人竹闊針混	640	其他墾地

600	灌木林	650	伐木跡地
611	天然草生地	700	道路
612	箭竹地	710	建築用地
613	牧草地	720	苗圃用地
620	茶園	730	水田
621	甘蔗地	740	防火線
622	蔬菜地	750	工礦用地
629	其他旱作地	760	土場用地
631	香蕉園	770	墓地
632	鳳梨園	780	鹽田
633	柑橘園	790	魚塭
634	桃.李.梅園	800	其他
635	蘋果.梨.水蜜桃園	900	裸露地
636	檳榔園	930	水面(河床、溪流、水庫、池塘)

地形資料

(18) 海拔高：

從像片基本圖上讀出樣點的海拔高並填寫四位數，不足位數者左邊補“0”。例如海拔公尺則填“0100”。

(19) 坡度級：

由該照片樣點在像片基本圖上的位置之等高線，用坡度表來直接量測。其代號如下：

代 號	坡度級	坡 度	代 號	坡度級	坡 度
1	平：	0-5度	4	急：	26-35度
2	緩：	6-15度	5	險：	36-45度
3	中：	16-25度	6	絕：	>46度

(20) 坡向：

指照片樣點在像片基本圖上之大地形的坡向，共分為個方位，其代號如下：

代號	方位	代號	方位
0	平地	5	南
1	北	6	西南
2	東北	7	西
3	東	8	西北
4	東南		

國立中興大學森林學研究所七十九學年度第二學期專題討論\

\2

植群分析樣區大小與形狀之探討\

A Research of Vegetation Analysis\

Quadrats Size and Shape\

\10

日期：80年 3月14日\

研究生：羅 南 璋\

\3

摘要\

\0

植群分析往往由於調查面積過大，若不以取樣分析法加以調查分析，實難以完成，且時間及物力之花費頗大。適當的取樣不但可以充分的反映出植物社會特徵，也可以節省時間及物力。本省生態學者於調查時多主觀的決定樣區大小，本文將國外生態學者常用，有關樣區之決定之方法，於此提出加以介紹。\\

\3

一、前言\

\0

生態學之領域包羅萬象，其中研究動植物之群聚現象，並推測影響其之環境因子及其歷史因素者，稱為生物社會生態學，或簡稱社會生態學，而以植物為對象者，特稱為植群生態學。\\

植群生態學之主要研究內容，乃在探討植物種類或植物社會與環境因子之關係，不同的植物社會，可以反應出環境因子之變異，故生態學之研究莫不從植物社會之分類著手。關於植物社會調查方面，往往因為面積太大，觀測不易，為節省時間及勞力起見，故常以分析取樣法加以調查。於取樣調查時，樣區面積的大小及形狀，常會對分析結果有所影響，且有效之取樣，可使人力物力做最佳之利用。用在生態調查方法中，關於樣區面積大小之決定，一般學者多依”種數—面積曲線 (Species-area curve)”來加以決定。本省有關植群分析之研究頗多，然均未調查種數與面積之關係，對於樣區面積之決定，多以主觀的 10x10m 大小來加以調查。\\

然此樣區面積究竟是否能充分反映植物社會之特性則不得而知。今就樣區之大小及形狀等問題於此提出並加以探討。\\

\3

二、最小樣區面積之決定\

\0

在取樣調查時，最感困難且首先要決定者，即為應調查幾個樣區及樣區的面積應採多大。調查樣區所需之最小面積視植物社會種類而異，一般在溫帶地區的植群

，可參考下列經驗所得來的面積大小來決定。

4

森林（喬木層）	200-500m
森林（僅林床層）	50-200m
乾草地	50-100m
低矮灌叢	10- 25m
牧草地	10- 25m
施肥之草原	5- 10m
農耕之雜草地	25-100m
鮮苔植物社會	1- 4m
地衣植物社會	0.1- 1m

0

實際僅有當植物社會植物社會組成較為均質，且不為零星分布時，方可能去決定調查所需的最小面積。除上述的經驗法則外，調查所需之最小樣區面積的決定，在定性的分析上，一般可採用種數—面積區線法加以決定。

早期歐洲之生態學者，重視定性之原

則，故對於樣區面積之決定，常參考樣區中植物種數與樣區面積之關係。在自然植物社會中，若增加樣區之面積，則植物種數亦必增加然增加之速率，視林分地點或地理區域而異。植相(Flora)豐富之地區，種數增加速率較快。反之，則增加速率緩慢。在同一地理區域之某一林分中，其增加速度，起初很快，後來則趨於緩慢。若以種數及面積代表直角坐標的兩軸，則由此畫出之曲線，稱為種數--面積曲線(Species-area curve)，可做為樣區大小之決定之參考。此曲線之斜率起初很大，後來漸趨緩和，並有漸趨水平之趨勢。由曲線上主觀決定一點，使植物種數增加率與面積增加率之比(即曲線斜率)，達到某一水準，此時之面積大小，即所謂最小面積(Minimal area)於此以下列的例子加以說明

種數—面積曲線法

此一方法，是由一個非常小之面積，譬如0.5m×0.5m(0.25m² 開始)，如圖 1，然後記錄所有

出現於此一樣區中植物種類。再將樣區面積放大兩倍，調查其所出現之新的植物種類，

以後亦逐次將樣區之面積兩倍、兩倍予以放大，然後依次調查新增加的植物種類，記錄之結

果如表1，一直到新的植物種類罕再增加為止。

然後以調查的樣區面積為x軸，所出現的植物種類為y軸，畫成種數—面積曲線，而所需之最小

面積，即為此一曲線斜率變小，曲線近乎水平之點上的面積。

5

圖 1、用來決定最小樣區面積的巢狀樣區系統。樣區號碼為連續的，由 1 開始，

且均包含前面的樣區。因此奇數號碼的樣區為正方形者；而偶數號碼之樣區則為長方形者。

例如表 1 即為依圖 1 之巢狀樣區系統，調查德國西北部一處草原所出現之植物種類，所得之結果：

Cain(1938)指出此種圖形所代表之最小面積，易受縱座標與橫座標之影響。若將縱座標對橫座標的比例予以縮小，則表面看起來，所須之最小面積也跟著縮小。為了克服此種缺點，Cain 便為最小面積圖定了一個不會影響 y/x 比值的標準。Cain 氏的建議，當樣區面積增加 10%，而種數也僅增加 10%，這時沿著取線上的點所對下來的面積及為所得須的最小面積。例如圖 2，10% 為 6.4m 種數的 10% 為 2.6，將此點描於圖上，然後將此點與原點連接起來。其實更精確的畫法，可將原點與 x 和 y 的 100% 之值連接起來即可(如圖二的 a)。然後再畫一條平行於此線的種數一面積曲線的切線，切點所對的 x 軸座標值，此值即為最小面積。對最小面積比較保守的估計，為當面積增加 10%，而種數僅增加 5% 時所對的面積。根據上述之方法求得本例所需的最小面積分別為：

種數增加 10% 時	8m
種數增加 5% 時	10m

由於此法所求得的最小面積，僅為近似值，因此一般均採四捨五入法求到平方公尺為止。

三、樣區的形狀

面積樣品法因須於林地內劃出一定面積之林地，以調查植物種類及數量，故樣區之形狀也有列入考慮之必要。一般生態學者於調查時常採用的樣區形狀有下列幾種。

a. 等徑樣區\

如方形區、圓形區均屬之。方形樣區每邊均相等，設置較為方便，可在任何植群中取樣。圓形樣區的設置更為簡便，只須要設立中心點，以一定之半徑圍繞一圈即可但僅適用於構造較低矮之植群，在高大森林中，必受樹幹之阻礙，無法順利拉出一圓形區。 \

b. 長方形樣區\

當樣區寬度與長度不等時，即成長方形樣區(Rectangular plot)\

C. 帶截樣區\

長方形樣區若長度遠大於寬度時，樣區呈長帶狀，即稱為帶狀橫截樣區 (Belt transect plot) 長方形樣區及帶狀橫截樣區常用於推移帶(ecotone)之取樣，因推移帶之植群變化，常呈梯度，若樣區之長軸橫跨其變異方向，則所得之資料變異最小，較具代表性。 \

d. 點狀樣區\

面積樣品法之極端縮小，到沒有長度及寬度，僅有點的位置者，稱為點狀樣品法(Point sample method)。其方法略述如下：\

在林分內隨機設置一條測線，由線上之基點起，每隔30cm為一測點，以十測點為

一小段，各小段間隔為6m共取十小段，調查個小段內測點所接觸到的植物種類及

點數目，分別列表計載，以供植物介量計算之用。(謝萬權等，1977，1978)\

e. 直線橫截樣區\

將帶狀樣區之寬度縮小為零，則樣區呈一直線，稱為直線橫截樣區(Line transect)。此種樣區指有長度沒有寬度，適用於調查構造較低的植群被覆，或森林下層之植物，植物之數量，由截取線上之長度表示之，可做為優勢度之指示，而在所有線上出現之次數百分率可計為頻度。但本法不能計算株數或密度，其在測線上之截取物，為枝葉之覆蓋，故主要對象為灌木或草本。本法又稱為直線截取法。謝萬權等(1977)曾用此法調查泡桐林下之植物被覆並與其它方法比較之。 \

\3

四、樣區大小及形狀之影響\

\0

樣區的大小對於恆存度、頻度和部分依賴於頻度的群聚度指數，有重大的影響，但對優勢度和密度不發生影響。 \

取樣分析時，樣區面積過大時會造成植物出現之頻度增加，然面積不足時，將會無法充分的顯現出植物社會之特徵。一般而言，若調查之目的地偏重於定性時，面積大小之取決，可依種數面積曲線來加以決定。若偏重於定量時，樣區之大小則可依下列幾個原則加以決定 \

1. 所測植物體的大小及個體間的距離 \

2.樣區應包括此植物社會中大多數的種\

3.盡量避免樣區落於單一種的聚生簇叢中\

至於何種形狀的樣區較為適用，關於此一問題，Bormann(1953)曾設立集落樣區，用變方分析法加以研究，其結論謂長方形樣區較方形樣區適用，國外許多學者也作過許多不同形狀樣區之比較，幾乎所有的學者均同意，就相同面積而言，長方形樣區較等徑樣區適用。正如Clapham研究指出，長度為寬度16倍的樣區，比他

試驗過同面積長寬比較小的樣區更有效率。欲解釋此一現象並不困難，因為代表每個種的個體，傾向於聚生成等徑的簇叢，而一個長形的樣區，具有一次就能截取數個簇叢部分的高或然率，而不完全落於一個簇叢範圍內，反之等徑樣區可能落於一個簇叢內，或兩個簇叢間，至使無法充分反映出此植物社會之特性。\\

\\3

五、結論\\

\\0

種數及面積之關係不僅為植群分析有效之工具，亦為各群落特徵之一，然研究此種

關係須依賴精密之取樣，耗費時間及勞力頗大，致使本省多數生態學者並未普遍採用，不過往後之生態調查，實有將其列入須調查項目之列。\\

在進行大面積生態調查時，若能事先測知樣區所需之最小面積，研究者將可更有效

率的進行調查工作。\\

生態調查主要有兩大目標，一為定性(qualitative)，一為定量(quantitative)

，而兩者受樣區形狀及大小之影響有所不同，樣區之最小面積之決定亦不相同，故在作生態調查時，首先要先確定調查目的為何，再進一步定出最小面積及樣區形狀。\\

\\3

六、參考資料\\

\\0

1.黃增泉、陳秀琴 1968 植物社會調查法之比較 臺大實驗林研究報告65號\\

2.謝萬權、林文鎮、蔡進來 1977 以不同取樣方法調查泡桐林下植被之比較 林學季刊 10(

1):11-18\\

3.蘇鴻傑 1976 臺灣北部烏來一小集水區闊葉樹林群落生態之研究(一) 臺大實驗林研究報

告118:183-199\\

- 4.蘇鴻傑 1977 臺灣北部烏來一小集水區闊葉樹林群落生態之研究(二) 地形與樹木分部型
式及其取樣方法之關係 臺大實驗林研究報告119:201-215\
- 5.蘇鴻傑 1977 臺灣北部烏來一小集水區闊葉樹林群落生態之研究(三) 植物種數與面積關係之分析 臺大實驗林研究報告120:171-184\
- 6.劉棠瑞、蘇鴻傑 1983 森林植物生態學 臺灣商務印書館發行 p.172-195\
- 7.Kershaw K.A. 1964 Quantitative and dynamic ecology. 2nd ed. American Elsevier Pub. Co. , New York. p.13\
- 8.Daubenmire R. 1968 Plant communities: A text book of plant synecology. Harper & Row. Inc. , New York.(陳慶誠 譯 1982 新華書店北京發行所發行) p.92-108\
- 9.Ellenberg H. 1974 Aims and methods of vegetation ecology. John Wiley & Sons. , New York. p.47-51\

- 一、題目：資源調查樣本數決定之探討
- 二、摘要
- 三、前言
- 四、前人研究

樣本數的大小

取樣若不預先求出實際所需要之樣本大小，實乃不智之舉。一般所稱之樣本愈多愈善的取樣原則，僅在能免費獲得精確測定值之情況下成立。影響精密度之因子為族群變異數 σ 及以取樣單位數目所表示之樣本大小 n ，前者隨族群之結構狀態及取樣單位之劃分而各異，後者常受調查經費之限制。（楊1980）

樣本大小可一根據族群變異數及目標精密度作如下之推算：

$$\frac{\sigma}{\mu} \times \frac{N-n}{n} = t \times \frac{E}{\mu}$$

設 $\frac{\sigma}{\mu}$ 為林分變異係數，將其代入上式之後，再兩側平方則得：

$$t^2 (CV)^2 \left(\frac{N-n}{n} \right)^2 \leq E^2$$

由是求得：

$$n \geq \frac{N \times t^2 \times (CV)^2}{N + t^2 \times (CV)^2}$$

式中 n ：樣本大小

N ：取樣單位個數

E ：精密度

CV ：林分變異係數

t ：依之自由度由分佈表中查出之值

所以族群之變異數 σ 或變異係數 CV ，為表示族群結構變異情形之母數。根據經驗或預備調查樣本獲知關於族群均數及變異數之資訊時，可從算出之經費及樣本大小得到所要之精密度。然而假如變異數推算值與實際相差甚大時，則所算出之樣本大小僅供裝飾點綴之用，毫無應用價值。（楊1980）

樣本數決定因取樣範圍、林分變異情形、精密度、調查時間及經費等限制。所以取樣時可依不同需要而決定採用何種公式來決定樣本大小。如下列公式：

$$(1) n \geq \frac{t^2 \times (CV)^2}{1 + t^2 \times (CV)^2}$$

N ：全部樣本數

$(2) n \geq \frac{(E)^2}{N \times t^2 \times s^2}$ $(3) n \geq \frac{t^2 \times s^2}{(E)^2}$ $(4) S = \frac{s^2}{n} \left(\frac{N-n}{N} \right)$ $(5) n \geq \frac{C - C_0}{C_1}$	<p>n：樣本數</p> <p>E：容許樣本誤差百分比(精密度)</p> <p>t：n－1之自由度(查t分佈表)</p> <p>s²：樣本的標準誤差</p> <p>cv：林分變異情形</p> <p>S\bar{x}^2：平均標準誤差</p> <p>c：預定調查經費</p> <p>C₀：一般開支</p> <p>C₁：每樣本花費</p>
--	---

如果只考慮經費則可採用(5)式，否則皆需考慮林分變異的情形或標準誤差等因子，其中(5)式並未考慮取樣的容許誤差，而(1)(2)(3)(4)式則均有考慮，並以(2)式之考慮可含概其他各式，所以較受廣用。

五、研究範圍

新化林場位於北緯°至°′″，東經°′″至°′″之間，屬台南縣新化鎮礁坑里。為一丘陵地帶，其地質屬粘板岩，土壤為砂質壤土，最高嶺之觀音山僅公尺高，為一水源涵養土砂捍止保安林，林場內共分十個林班、面積.5835公頃，其中包括國有林地面積筆共.9226公頃，保育林、私有林地面積筆共.6609公頃。如圖六：新化林場林班圖。圖七：新化林場林型圖（資料來源：國立中興大學實驗林管理處新化實驗林場公私有地清查名冊，）

黃色：林班線 藍色：林道
圖六、新化林場林班圖（陳 1993）

綠色：大葉桃花心木	白色：龍眼
黃色：大葉桃花心木、柚木混交	咖啡色：大葉桃花心木、相思
深綠：柚木	樹混交
淺藍：大葉桃花心木、雜木混交	藍色波紋：麻竹
淡綠：雜木	白斜綠：相思樹

樹種	林型別 面積 (ha)	取樣 單位 (N)	實設 樣區數	林分變異 (CV)	應設樣區 數 $t(\alpha)$ 0.3	應設樣區 數 $t(\alpha)$ 0.2	應設樣區 數 $t(\alpha)$ 0.1
大葉 桃花心木	116.17	2323	33	0.7344	7	21	137
麻六甲 合歡	35.12	702	5	0.1479	1	1	8
龍眼	52.87	1057	6	0.1356	1	1	7
芒果	1.20	22	3	0.8376	8	13	21
柚木	14.17	283	3	0.3306	2	7	50
苦楝	3.24	65	3	0.3414	2	20	32
桉樹	0.79	16	3	0.1753	1	2	8
相思樹	33.14	663	3	0.2640	2	5	36
樟樹	6.13	123	3	0.5823	6	12	74
合計	262.75	5254	62		30	82	373

由表十三得知三種不同容許誤差範圍下信賴區間之統計值取樣個數，分別為誤差率% 精密度% 時應設個樣區，誤差率%精密度% 時應設個樣區，誤差率%精密度% 時應設373個樣區。並且由誤差率控制在%精密度%時，大部份樹種的應設樣區都可以達到應設樣區的數量，但是芒果和樟樹實設樣區數少於應設樣區數，主要原因在於林分變異較大，而新化林場由於芒果種植之行株距較一般樹種為寬（為求果實），且為闊葉樹幹形分歧多，所以在本研究中林分變異以材積時，亦造成此現象。至於樟樹因為新化林場種植不久，約為5年生（造林台帳資料），所以生長良莠不齊，是造成林分變異較大的原因。所以經營者可以依不同需求的精密度來設樣區。

（二）樣區材積的信賴區間

如表十四：樣區材積統計表所示，並經由以下之運算可得樣區材積的信賴區間。

$$\text{均數變異數 } S_{xh}^2 = \frac{S^2}{nh} \left(1 - \frac{nh}{N} \right)$$

$$\text{族群平均值 } x_{st} = \sum \frac{Nhx_h}{N} = \frac{26156.7696}{5254} = 4.9784 \quad (m^3/0.05ha)$$

$$\overline{Sx}_{st} = \frac{(\sum_H S2)}{N2 \cdot XH} = \frac{3249099.805}{52542} = 0.3431 \text{ (m3/0.05ha)}$$

$$95\% \text{信賴區間} = xst \pm \frac{t}{2} \cdot Sx_{st} = 4.9784 \pm 0.6862 \text{ (m3/0.05ha)}$$

$$99\% \text{信賴區間} = xst \pm \frac{t}{2} \cdot Sx_{st} = 4.9784 \pm 0.9126 \text{ (m3/0.05ha)}$$

$$\text{機差百分率 (Percent of standard error)} = \frac{\frac{t}{2} \cdot Sx_{st}}{xst} \times 100\% = 13.78\%$$

$$= \frac{2 \times 0.3431}{4.9784} \times 100\% = 13.78\%$$

得知機差百分率為.78%(顯著水準為5%)，稍微超過%(林、賴、伍 1979)
可知調查樣區數區，在材積推估時，仍嫌不夠。

表十四、樣區材積統計表

樹 種	面 積 (ha)	可設樣區數 (Nh)	實測樣區數 (nh)	每樣區平均材積 (xh)(m/0.05ha)	變異數 (S2)	均數變異數 (S2)
					H	XH
大葉桃花心木	116.17	2323	33	5.5527	16.6293	0.4967
麻 六 甲	35.12	702	5	6.4731	0.9166	0.1820
龍 眼	52.87	1057	6	2.8717	0.1516	0.0251

芒	果	1.20	22	3	2.6109	4.7825	1.3767
柚	木	14.17	283	3	4.1398	1.8731	0.6178
苦	楝	3.24	65	3	2.0772	0.5029	0.1599
桉	樹	0.79	16	3	1.0783	0.0357	0.0097
相	思 樹	33.14	663	3	6.2553	2.7271	0.9048
樟	樹	6.13	123	3	1.2155	0.5010	0.1629
合	計	.37	5254	62			

八、結論

以新化林場為例實施永久樣區的設置與主要林木資源的調查，藉以探討資源調查樣本數大小的決定，得知三種不同容許誤差範圍下信賴區間之統計值取樣個數，分別為誤差率% 精密度% 時應設個樣區，誤差率% 精密度% 時應設個樣區，誤差率% 精密度% 時應設373個樣區，以及各主要林型分別所應設之樣區個數。本次新化林場樣區材積在%信賴區間時，為 $.9784 \pm 0.6862$ (m³/0.05ha)，並且求得機差百分率為.78%（顯著水準為%），稍微超過%，可知調查樣區區，在樣區材積推估時的代表性仍嫌不夠，所以未來應再增設永久樣區。

九、引用文獻

1. 楊榮啟 1980 森林測計學 國立編譯館主編 頁167-316

森林調查法 Forest Survey Method

一、全林直接測計法:依林積推算方式不同分為:

- 1.立木材積表法 歸推算法
- 2.標準木法
 >比率推算法
- 3.標準地法
 - a.由每木調查結果 >推算林分材積之方法
 - b.每木調查在晚秋、初春時較適合
 - c.立木材積表法
用(a)僅包括胸高直徑之一獨立變數立木材積表上查出材積
(b)應用樹高曲線求樹高 >胸高直徑+ 樹高二個"獨立變數立木材積表"查材積
(c)胸高形數表查形數 >材積

二、取樣調法(Sampling method)

- 1.定義
欲知取樣對象事物(如林分或森林)之母集團(population)之構造,而於母族群之中抽取一部分調查材料,即樣本(sample)或樣區(plot),以此樣本來推測對象事物之某種狀態,即稱取樣調查法(sampling method, sample survey)
- 2.取樣調查(樣本調查)操作程序
 - (1)使用逢機方法從族群中選取樣本個體
 - (2)測定樣本個體的性態值
 - (3)從已知樣本值推論未知的族群值
3. 目的
 - (1)可避免人為選擇誤差,提高結果準確性
 - (2)可分a. 目的取樣法(purposive sampling),仍有人或物之偏性(Bias)
 b. 機率取樣法(probability sampling),多採用之。

(三)機率取樣法

1. 基於機率抽出的作法,而由全林中抽出之標木及標區。除非調查困難執行或其費用過分昂貴,總以合乎統計原理的取樣為明智之舉。
2. 與標準木、標準地不同台灣林相複雜,不均齊適用之,取樣法則

為提高精密度必須抽選較多樣區。而本二法可節省尋找及到達樣區的時間。

(四)森林測計實行取樣調查之理由

1. 調查受時間及經費的限制 諱祧d對象,而將節餘時間與經費增加調查項目及結果整理分析。
2. 無實行全林調查之必要 @業易改進,非取樣誤差機會減少,雖比全林測計之精密度差些,但在不需太高精密度時,可簡單又迅速求出所要的結果。
3. 調查目的相同欲使調查結果更豐富 p同時要調查林木材積>非林木價值,森林結構推移之資訊。
4. 根本不可能實施全面調查 p測計林木生長量可作樹幹解析&測定木材各項性質所作的破壞性調查時,非採取樣調查不可。

二、林林調查所使用之取樣設計可概括分類如下:

A. 逢機取樣(random sampling)(機率取樣)

1. 等機率取樣(Sampling with equal probabilities)

(1) 樣區法(sample plot)

(2) 樣木法(sample tree)

2. 不等機率取樣或變動機率取樣(sampling with varying probabilities)

(1) 列表取樣(list sampling)

(2) 3P取樣(3P sampling)

(3) 樣點法(point sampling; PPS sampling)

(4) 樣線法(ling sampling)

B. 非逢機取樣(non-random sampling)

1. 選擇取樣(selective sampling)

(1) 標準地法(purposive sample plot method)

(2) 標準木法(probestamm, mean tree)

2. 系統取樣(systematic sampling)

C. 二段取樣法(two-stage sampling)

.等機率取樣(Sampling with equal probabilities)

(1) 樣區法(sample plot)

D.f: 由全林或一地區選取能代表該林區或地區之較小面積,從事調查測計由小面積所得資訊 >再推算全林區或地區之值稱之.由標準地(probeflache)法演變而來,由測定者主觀意識而定;精密度則依測計者經驗而得,後應用統計學中之取樣觀念由大面積中逢機取部分實施測計,再推算全面積

。

Note: 各林木或樣木之機會相等,所以屬等機率(equal probability y)調查法。氏認為樣木各直徑級之分佈與該林分各直徑級林木之株數成比例,所以稱其為P.P.F(probability proportional to frequency)因各樣區面積固定,所以又稱固定樣區取樣法(fixed plot sampling)

(2)樣木法(sample tree)

應用逢機取樣法抽出作為樣本之樹木(與使用非逢機選擇法所得之標準木有別)特稱樣木。應用逢機取樣法可使用統計方法處理,能客觀詳定誤差大小。

.不等機率取樣或變動機率取樣(sampling with varying probabilities)

就是:調整群體的選取機率,使較大或較貴重之群體組成分子有較大之機會被選作樣本。

(1)列表取樣(list sampling)

取樣機率與表中所列項目之有關數量成比例。

(2)3P取樣(3P sampling)

說明:(1)所謂P是"Probability Proportional to Prediction

"也就是:使取樣機率與預測量成正比。

(a)3P取樣法主要用於森林內部林木售賣時推算材積材積及材價用。

(b)3P取樣為不歸還之逢機取樣

方法:(1)調查者進入林內檢視每木,估測其(DBH),並算出(DBH)²,以之表示,將這些估測資料列表記錄如下:

Xi		Xi	
薹	(DBH) (DBH) ²	薹	Ri (DBH) ²
1		1	
2		2	
3		3	
4		4	

()2

- (c)於逢機數字表中採不歸還逢機取樣,抽出個逢機數字(R_i),再將 $(DBH)^2 \geq R_i$ 則該株木為錄取的樣木,若 $(DBH)^2 < R_i$ 則摒棄之。
- (d)實測被錄取樣木之材積(Y_i),依下面公式推算林木之總材積(Y_a)

$$Y_a = \frac{(DBH)^2}{N} \left(\frac{Y_l}{X_l} + \frac{Y_m}{X_m} + \frac{Y_n}{X_n} \right)$$

N:表示被錄取樣木的株數

Y_l, Y_m, Y_n ____表第號,第號,第號的實測材積

X_l, X_m, X_n ____表第號,第號,第號的 $(DBH)^2$

(l,m,n____表被選取樣木的樹號)

(3)樣點法(point sampling; PPS sampling):Bitterlich point sampling

(a)P.P.S.就是"取樣機率與取樣單位大小成正比"

(b)森林測計學中的.P.S.具有廣義,狹義兩種意義:

廣義:與變動機率取樣的概念相符合。

狹義:則為森林取樣方法之一,單株樹木的選取機率與表示樹木可測因子之大小成比例。按照取樣機率與被測定個體之某一因子大小成比例的原理進行取樣。即不同林型之林分或不同大小之林木使用不同大小之樣區(如:2000株/ha 幼齡林 >取.05ha即可, 10株/ha 老齡林 >取0.2ha才可)測者立於樣點上,使用水平定角儀透視其周圍立木胸高直徑,計算大於定角透視線之立木株數。

各級之樣點立木株數

單位面積之林木

w數 =

樣點數

DBH basal area 之推算值

因為不需測樣區面積, 所以又稱無邊樣區取樣法(potless sampling)
因為區別其與固定樣區法,特將樣點取樣稱為P.P.S.(sampling
with probability proportional to size)選取機率與其大小成比例
之取樣法。

(4)樣線法(line sampling)

圖. 水平線取樣法林木之選擇

figure 1. Selection of Trees in Horizontal Line Sampling

Associated plot= $W_L=2R_L$ =function of D_L

=要測定之樣木 in tree sample

=夾邊林木 border line tree

=不要測定之林木 out tree

=林木最大距離 tree maximum distance

(m)=水平距離定乘數乘林木直徑(cm) horizontal distance multiplier x tree
diamete(cm)

=樣區寬度 Plot width

.非逢機取樣(non-random sampling)

.選擇取樣(selective sampling)

依測計者的主觀判斷選取樣號(但因為個人主觀成見使推算結果產生偏差,且無法求信賴度)

(1)標準地法(purposive sample plot method)

依標準地測定全區域內某種狀態(如材積生長量____),推算全林分的方法。

標準地:指一定區域森林中,選出具有代表該森林之某種狀態之特定小區域(selective sample plot)。

.受測定者主觀影響甚大,準確度依技術熟練而異。

.大面積森林中,覓定足以代表森林平均狀態的標準地屬困難。

.所以調查對象森林面積較少,且林相整齊或林木價低之林分可依此法測之。

$$\text{全林面積} = \frac{\text{標準地材價合計}}{\text{標準地面積合計}} \times \text{全林總面積}$$

(2)標準木法(probestamm, mean tree)

使用標準木推算森林材積的方法 >非隨機取樣或稱目的選擇 >無法推算精密度,統計

學立場不穩少用之。

即由標準木材積以林木株數或林分斷面積為媒介推算全林材積之方法。

如:Hartig法:全林林木,按直徑大小順序分成胸高斷面積合計相等或概略相等的直徑級,各級

分別決定標準木,全林材積推算。

.系統取樣(systematic sampling)亦稱機械取樣法(Mechanical sampling)

僅第一樣區係以隨機方法選出,其餘樣區則以預定之間隔距離,配置於全林區內。

優點:(1)呈規則間隔,尋找樣區可沿固定方位角及距離前進,減少取樣巡行時間,迅速,費用低

。

(2)樣木均勻分散族群中,更能代表族群。

(3)起點決定後,即按固定間隔選出,可一直繼續進行,雖無法預先知道族群大小時亦可用。

(4)測計者可同時完成繪製林型圖。

(5)系統分佈族群中,本法可用於固定樣區或樣帶法應用,亦可在.P.S.取樣之變動樣區的樣

點法,樣線法。

(6)樣區在全林區內分佈均勻,樣本平均值與族群真平均值間常具較小機差,通常其精密度常

不差於逢機取樣,用逢機取樣公式計算,一般高估。估值之標準誤差可以更正之。

.二段(副次)取樣法(two-stage sampling)

系統取樣+逢機取樣

大面積森林調查時,先將對象森林分成若干個大區別稱初級取樣單位 >逢機取之

再將一級(初級)取樣單位分成若干個樣區,然一級樣區先取再由之逢機取出二級單位,稱二次

(二級)取樣單位 >系統取之。

.森林生長調查所用之取樣法

連續取樣法(或稱重複取樣法)

.前次(第一次)調查時森林特性之數量。

.第二次調查時森林特性之數量。

.推算兩次間隔期間之變化或生長量。

.分層逢機取樣法

分層逢機取樣之二次理由

.求得各次級族群的推算值。

.減少推算整個族群的取樣誤差。

多段取樣在森林調查上應用之樣點:

.減少行走於各次級單位之時間。

.外業測定工作集中於少數幾個位置監督容,易能減少非取樣誤差。

三、取樣調查的統計分析

(1)簡單逢機取樣

逢機抽取:由含有個取樣單位族群中,抽出個取樣單位的所有可能的組合。各取樣單位皆有

同樣機率被選取機會

$$\bar{Y} = \frac{1}{n} \sum_{I=1}^n y_i$$

n

$$S^2 = 1/(n-1) \sum_{i=1} (y_i - \bar{y})^2$$

$$\bar{S^2_y} = S^2/n(N-n/N)$$

$\bar{S^2_y}$: 樣本均數的標差(或稱標準誤差)

S^2 : 樣本的變異數 S : 樣本標準差

n : 樣本含取樣的單位數目

N : 族群所含的全部取樣單位數目

$(N-n/N) \rightarrow 1$ 當無限族群時

自由度 機率%的值($tn\%$)

信賴區間: $\bar{Y} \pm \bar{y}$

誤差率 $E\% = (t \cdot \bar{y}/\bar{y}) \cdot 100\%$

(3) 雙重取樣(double sampling), 二相取樣(double-phase sampling)

主要變數 輔助變數

$$Y = f(X)$$

population n 單位之輔助變數 m 個單位主要變數

簡 逢 簡 逢

變數 X 變數 Y

$$\bar{Y}_{reg} = \bar{Y}_m + b(\bar{X}_n - \bar{X}_m)$$

(2)系統取樣調法

$$n = \frac{4c^2A}{e^2A + 4ac^2}$$

A:全面積11.50ha
e:誤差率%(精密度90%)
c:林分變係數20%
a:樣區面積0.05ha

樣區取樣率

$$\begin{aligned} P &= \frac{n}{N} \times 100 \\ &= \frac{(N - n)}{N} \times 100 \\ &= (15 - 0.05 / 11.5) \times 100 \\ &= 6.5\% \end{aligned}$$

樣區取樣誤差率

$$\begin{aligned} e &= 2c \sqrt{(A - a) / n} \times 100 \\ &= 2 \times 0.02 \sqrt{(11.5 - 0.05) / 16} \times 100 = 9.6\% \end{aligned}$$

樣區間隔

$$d = \sqrt{\frac{A}{n}} \times 100 = \sqrt{11.5 / 16} \times 100 = 84.8 \text{ m}$$

調查結果

樣區號	株數	林積(y)	x ²	xy	y ²
1	45	15.66			
2	—	12.27			
3	—	—			
—		—			
—		—			
—		—			
16			2	y	2
計	1035				

立木株數之推估(x)

1.每一樣區之平均株數

$$\bar{X} = \frac{\sum x}{n} = 1035/16 = 64.69$$

2.平均值之變異(分散)

$$V_x = \frac{\sum (x - \bar{x})^2}{n(n-1)} = \frac{N - \frac{(\sum x)^2}{N}}{n(n-1)}$$

$$= \frac{69.735 - 64.69 \times 1.035}{16(16-1)} = \frac{230 - 16}{230}$$

$$= 11.59 \div 0.93 = 10.78$$

3.平均值之標準偏差

$$S_x = \sqrt{\bar{x}^2 - 10.78} = 3.28$$

4.全株數之推估(精密度90%)

$$\begin{aligned}Tx &= A/a[\bar{x} - \bar{x}] \\ &= 11.50/0.05[64.69 - 1.753 - 3.28] \\ &= 230 - (64.69 - 5.75) \\ &= 13.556 - 14.311\end{aligned}$$

5.標準誤差率

$$e = \bar{s}/\bar{x} \times 100 = 3.28/64.69 \times 100 = 5.10\%$$

6.林分變異係數

$$\begin{aligned}Cx &= ax/x \times 100 \\ &= \frac{\sum (x - \bar{x})^2 / (n-1) \times (N-1) / N}{\bar{x}} \times 100 \\ &= \frac{172.41}{64.69} \times 100 \\ &= 13.13/64.9 \times 100 = 20.30\%\end{aligned}$$

林分材積之推估(y)

1.每一樣區之平均材積

$$\bar{Y} = \bar{y} = 293.604/16 = 18.350$$

2.平均值之變異

$$V_y = \frac{\sum (y - \bar{y})^2}{n(n-1)} = \frac{N-1}{N} \cdot \frac{\sum y^2 - (\sum y)^2/n}{N-1}$$

$$= \frac{5662.849 - 18.350^2 \cdot 16}{16(16-1)} = \frac{293.6204}{230}$$

$$= 1.147 \cdot 0.93 = 1.069$$

3.平均值之標準偏差

$$S_y = \sqrt{V_y} = 1.067 = 1.03$$

4.全林材積之推估(精密度%) _時間、能力、財力

$$T_y = A/a [\bar{y} + t \cdot S_y]$$

$$= 11.50 / 0.05 (18.350 \pm 1.735 \cdot 1.03)$$

$$= 330 \pm 18.26$$

$$= 3.805 \sim 4.635 \text{ m}^3$$

$$(4220 \pm 415 \text{ m}^3)$$

5.標準誤差

$$C_y = \frac{17.06}{100} = 17.06\%$$

y 18.35
=22.51%

地理資訊系統於分層取樣調查設計與分析上之應用
Geographic Information System Used to Stratified Sampling
Design and Analysis

1. 地理資訊系統之空間圖層資料
2. 地理資訊系統之空間分析套疊分析
3. 分層取樣調查設計
 事前分層
4. 取樣調查資料分析與展示
 事後分層

(一) 分層取樣調查設計

以往的資源調查雖名為雙重取樣，但其實只是利用簡單逢機 (sample random) 或系統取樣 (systematic field sampling) 以求得航空照片樣點與地面樣區，再分別由航空照片判釋材積和地面調查求取樣區材積以為蓄積推測之依據而已，並沒有將航測資料與地面樣區資料加以整合，以改進材積推估能力。

另外有關分層問題，假如分層 (stratification) 成功則樣區或樣點可以充分代表不同族群，若以相同的經費來加以比較，則分層取樣將可獲得較精確的材積。然分層取樣可分為事前分層與事後分層，亦即調查前、後依不同林分性質進行分層，以利調查或分析，台灣三次資源調查皆係採事後分層。

分層可由圖或像片加以取樣，然一般由航空照片來得多，整合照片樣區與地面樣區為分層取樣設計，為較省錢的工作。然據研究由蓄積級分層是無法減少推算生長收穫的變異 (variance)，即依蓄積級分層並無法改善蓄積推算值 (Maclean, 1972)；其精確性 (precision) 則完全依照取樣數量來決定。

1、分層選點

- 森林分層可依 (1) 地形 (topographical feature)
(2) 地位級 (site classes)
(3) 坡度 (aspect)
(4) 坡向 (slope)
(5) 林型 (forest type)
(6) 密度級 (density classes)
(7) 材積 (volume)
(8) 樹高
(9) 年齡
(10) 海拔高 (altitude)
(11) 事業區 (working circle)
(12) 集水區 (watershed)

分層之基礎盡可能以推估時相同性態值為之，目前改進以地理資訊系統之不同圖層 (map layers) 重疊以獲所選定之區劃範圍。

2、分層取樣

分層(stratification)簡單講即為等級的使用，應用於取樣即為在同一等級裡其個體皆相類似(級內的變異很小)，在其他等級裡則不同。若應用分層於森林資源調查上，則可以以較少的勞力、經費來調查樣本，而達到相同的正確度(accuracy)。

適當的使用分層將可以提高取樣的效率(efficient sampling)，何謂適當，可以分成三項來加以說明：

- (1) 各定義層(defining strata)一定要獨立取樣以獲各層級的平均值。各層要獨立取樣，意即各層級的分布必須於個體被選擇計算各層級平均值前即已知。在森林資源調查上有許多原則可供分層級，如地理因子(geographical factor)、林型(forest type)或林分狀況(stand condition)等。這些因子可由以往的圖籍、航空照片判釋或事先的地面調查等事先查知。目前最有效的過程為利用高品質的航空照片、高技術的像片判釋和遙感探測技術等以便獲得地理資訊系統內所需要的高品質資料來加以研判。
- (2) 各層級的平均值必須確實不同。此平均值係指樣本的平均值(arithmetic average or mean)與離均(平方)差(average squared deviation from mean or variance)、變異(variance)或均數的變異數或標準差(標準誤差)，各層級間的差異愈大，則分層取樣之效率愈較簡單隨機取樣來得高。
- (3) 各層級的觀測值分布一定要適當。決定適當與否有二法，一為比例分派(proportional allocation)，二為最適分派(optional allocation)。
 - A. 各層級內的變異(variance)皆相同，若這樣則依各層級的權重的比例予以分派，於資源調查中，各層級的面積是分派的權重依據。
 - B. 若層級的變異數不相同，比例分派仍然可用，但最適分派(optimum or Neyman allocation)則較有效率(即以較少的成本獲得想達成的正確度)。最適分派意為觀測值與(各級面積×標準偏差)乘積呈比例分佈(distributed proportional to the product of class area & standard deviation)。(註：標準偏差standard deviation = $\sqrt{\text{variance}}$)
 - C. 若觀測值其各層級的變異既不相同且不呈比例，則分層取樣對此族群沒有效率。

由以上所提，一般地理因子(geographic commits)如鄉鎮、集水區可

能無法由分層取樣獲益，林型會好些。

如何分層級，可由許多原則決定，但最有效的原則還是與欲推估事物有密切關連者。

林木資源調查(timber cruising)的目的為推算材積。在這種情況最有效的分層標準為減少同層級內材積變異，或增加層級間級平均值的差異。

(二) 永久樣區之取樣規劃實務

1、永久樣區之取樣規劃

以新化林場地理資訊系統資料庫所建立之圖層(馮等 1993, 陳 1993), 配合地理資訊系統套疊(overlay)、選擇(select)之功能, 將林地內具有不同地理條件之地點加以標示, 以不同之樹種、林齡、地理座標、海拔高土壤類型、坡度、坡向做為分層取樣, 規劃永久樣區之依據。規劃步驟如下所列:

(1) 資料蒐集

蒐集造林台帳、像片基本圖、航空照片、歷年研究文獻及數位地形模型(Digital Terrain Model, DTM)。

(2) 土地利用型圖(含林型圖)及林業主題圖之彙整

依陳(1993)所建之新化林場地理資訊圖檔(林場界、林班界、林道、水系、土地利用型、等高線、坡度、坡向、立體圖及土壤圖)為基礎, 轉換成相同網格大小之網格式圖檔, 土地利用型圖依不同屬性及其林型代號(tag), 查詢出單一土地利用型(含林型)。

(3) 圖形套疊及樣區選點

以林班界圖、土壤圖、坡度圖、坡向圖, 配合已建立之資料庫圖層分別與不同的樹種(人工林)或林型(天然林)圖, 加以套疊, 產生不同地理因子與土壤類型之林型圖, 再分別對應造林台帳決定林齡。至此, 即可於圖面上將相同之林型, 不同坡度、坡向和土壤類型之各種組合區域標示出來, 並於各區域中以網格板逢機決定樣區位置。

(4) 尋找樣區、推算調查費用

套疊林道圖、水系圖林班界圖即可獲知樣區之座標以及到達樣區所需路徑、地理狀況, 以便計算樣區調查時間、費用, 更可配合GPS或航照片、基本圖, 尋找出樣區正確位置。

永久樣區坡度、坡向分層之檢定:

以Bertalanffy模式預測各永久樣區之基準林齡(14 年生)胸高

直徑生長量，再以進行非均衡性變方分析，結果顯示：坡度、坡向之群間皆具有顯著差異，但坡度、坡向間之交感效應即未檢定出變異，如表一。故可知永久樣區之設立上，分層取樣之考量仍有其具體意義。

表一 不同坡度、坡向樣本變方分析表

變因	DF	SS	MS	F Value
坡度	2	25.441956	12.720978	* 6.17
坡向	3	21.403898	7.134633	* 3.46
交感效應	6	20.492197	3.415366	1.66

*表差異顯著水準% (MSE=2.061746)

*DF 為分群後之自由度

以新化林場而言，樹種組成包括了大葉桃花心木、柚木、竹、龍眼、麻六甲合歡、相思樹、樟樹、黑板樹、芒果、桉樹、苦楝等，其中以大葉桃花心木最為重要，在永久樣區之設定上實應優先考量；若欲配合林場地理資訊系統之建立，則其地理因子更突顯出在取樣上的重要性。新化林場海拔高度在公尺至152公尺間，變化不大，坡度、坡向在本研究中證實對林木造成生長上的差異，可提供今後永久樣區設置時之參考。

在實際作業上，由於坡度、坡向二種地理因子無法在分層取樣中做細度的分級，本實驗坡向以方位角表示且 $^{\circ}$ 為一等級，若依Whittaker(1960)提出之方位與水分梯度之關係，如圖一；則亦可由水分梯度來探討其對林木生長之影響。

新化林場之地勢較平坦，坡度分級以度為一級。

<hr/>									
NE		ENE		E	ESE	SE	SSE	S	SSW
NNE		N		NNW	NW	WNW	W	WSW	SW
<hr/>									
1	2	3	4	5	6	7	8	9	10

圖一 根據方位之水分梯度 (Wittaker 1960)

(三) 永久樣區之取樣規劃及設立

利用新化林場已建立之森林經營資料庫(陳1993)，做為大葉桃花心木永久樣區取樣規劃之依據，林型圖、土壤圖、坡度圖、坡向圖之套疊，可得不同坡度級、坡向級、土壤類型之分層資料，提供樣區設置時之參考。舉例而言，坡度5-15度、坡向為北向，土壤類型為壤土之大葉桃花心木林型，其可供樣區

設立之範圍，可經由各相關圖層之交集分析產生，再經由林班區界、林道圖之疊合，提供樣區之位置(如圖九)，並達成以坡度、坡向、土壤類型為分層依據之目的。因新化林場之坡度區分為 6 級，坡向區分為 8 級(依據第三次資源調查)，同時包含 5 種土壤類型，疊合工作極為繁瑣，應考慮各個分級間之合併或設計輔助軟體，以節省作業時間。

4. 取樣調查資料分析與展示

事後分層

(二) 方法

- 1、資料處理—將文字檔轉入III和LOTUS 以方便整理。

將圖檔建入地理資訊系統中處理、展示。

- 2、資料分析—將地面樣區調查之單株林木資料進行林分結構與林分蓄積分析、並將結果累彙送至林分資料檔。

(1) 樹種組成

各樣區由航空照片判釋及地面植生踏勘所判別繪製之林型為植群分類並進行各樣區內樣木的樹種種類分析以得樹種組成依性質將樣區歸屬為主要樹種林型。

整理主要土地利用型內所含主要樹種林型樣區之百分比天然林之森林植群之組成、林分結構、林分蓄積是受生育環境影響。

樹種組成—用相對頻度、相對密度、相對優勢度之和表示之重要值。

重要值指數(Importance Value Index, IVI)

- A. 為學派改善指數(Density-Frequency-Dominance Index) 密度、頻度、優勢度(或覆蓋度)絕對值，指數大小沒有一定範圍，且三介量單位不一總和以代表的不合理現象。
- B. IVI為相對密度+相對頻度+相對優勢度，每個介量最大值為，植物重要值之總合恆為，若一植物社會值由一植物組成，則其重要值為300。
- C. 重要值可顯示某植物於該林型或林分中相對重要性，且兼顧密度、頻度及優勢度之綜合特性(劉、蘇1989)。

$$\text{相對頻度}\% = \frac{\text{某植物之頻度} \times 100}{\text{林分中所有植物頻度之總和}} \dots\dots\dots (1)$$

$$\begin{aligned} &\text{相對密度(relative density)\%} \\ &\quad \text{某一植物之株數} \\ &= \frac{\quad}{\quad} \times 100 \dots\dots\dots (2) \end{aligned}$$

所有樣區內全部植物之株數

$$\text{相對優勢度 (relative dominance) \%} = \frac{\text{某一植物之優勢度 (或覆蓋度)} \times 100}{\text{所有樹種優勢度之總和 (或覆蓋度)}} \dots\dots\dots (3)$$

$$\text{重要值指數 (Importance Value Index, IVI)} = \text{相對頻度} + \text{相對密度} + \text{相對優勢度} \dots\dots\dots (4)$$

(2) 林分結構

林分結構是指林地、樹種、林木各種性態值大小分布，為樹種生長習性，環境條件和經營實施的綜合結果 (Husch, and Beer 1972)。天然林的林分結構可由林木的性態值如胸高直徑、斷面積、樹高、材積大小的分布加以說明。所以直徑分布是林分結構中的一種表示方法，乃因胸高直徑資料最易獲得，故常被使用，而直徑分布以機率密度函數 (如 (5) 式) 描述得最理想 (馮1988)。

A. 各樣區內樣木的林分結構以機率密度函數 (x, θ) 來配合直徑分布，求算出的三個母數 a, b, c 之最佳推估值 $(a, b, c,)$ 。

$$f(x, \theta) = (c/b) [(x-a)/b]^{c-1} \exp[-(x-a)/b] C \dots\dots\dots (5)$$

式中的 $a \geq 0, b > 0, c > 0$

x 為胸高直徑

\exp 為指數

(3) 建立分析推估材積之模式系統推估林分蓄積

林分蓄積可由林分內單株林木材積累積，或直徑分布法 (diameter distribution method) 由各直徑級之材積累計，或全林分生長模式 (whole stand method) 由全林分材積與樹種、林齡、地位、密度、處理間的迴歸式推估而得。由於天然林林木種類複雜，林齡不易獲得，所以不加以考慮。環境因子對植生組成、結構、蓄積之影響，有層次且次序相關性存在，以地形位置 (topographic position) 及方位 (aspect) 為最上層因子，可解釋大部分植生之地域變異，稱為環境層級觀念 (Kellman 1980)。所以以坡度、坡向、海拔高等地理因子代表地位。密度則以單位面積之林木株樹表之。

A. 整理各重要樹種之單株林木材積式，樹高曲線式。

B. 利用單株林木材積式推估累計樣區內之林木材積 (SUV)。

以樣區面積除之，以獲得單位面積之林木蓄積 (SUV/ha)

$$SUV/ha = \sum V_i / ha \quad V_i = f(D, H) \dots \dots \dots (6)$$

式中的 V_i : 株林木材積
 D : 胸高直徑
 H : 樹高

C. 利用直徑分布法求算樣區之林木材積 (DDM)。

以樣區面積除之，以獲得單位面積之林木蓄積 (DDV/ha)

$$DDM/ha = Nt \int (x) f(x, \theta) \dots \dots \dots (7)$$

式中的 Nt 為樣區內林木株數換算之單位面積之林木株數
 $g(x)$ 為胸高直徑的函數式如 $H=f(D)$ $V_i=f(D, H)$
 $f(x, \theta)$: 的母數 θ — a, b, c

D. 將樣區內樣木所求算之 $main\ sps$, a , b , c , Nt/ha , SUV/ha , DDV/ha 轉入樣區林分檔中。

E. 繪製林分結構母數 (a, b, c) 與材積母數 (SUV/ha , $/ha$) 與坡度 ($slope$)、坡向 ($aspect$)、海拔高 (AH)、林齡 (age) 之關係圖由此判斷林分結構、材積與環境、林齡之間的關係為直線或指數、倒數。

F. 建立各 SPS 土地利用型與林型、林分結構、材積與地理因子，建立樹種空間性材積推估模式。

建立林分結構、林分蓄積分析模式系統架構。

各林型 $SUV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $DDV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $\theta = f(\text{坡度、坡向、海拔高})$

G. 將土地利用型歸併為種天然林林型及包含人工林、建地等其他土地利用型。依第三次資源調查所繪製之林型圖算出各種天然林涵蓋的面積以及顯示其所在位置範圍之天然林林型圖。

H. 利用 (c) (d) 所算出之單位面積之林木蓄積 (SUV/ha) 與 (DDV/ha)，配合 (G) 算出之各種天然林涵蓋的面積，可以以比率推算 ($ratio\ estimation$) 出各種天然林型之林木蓄積。

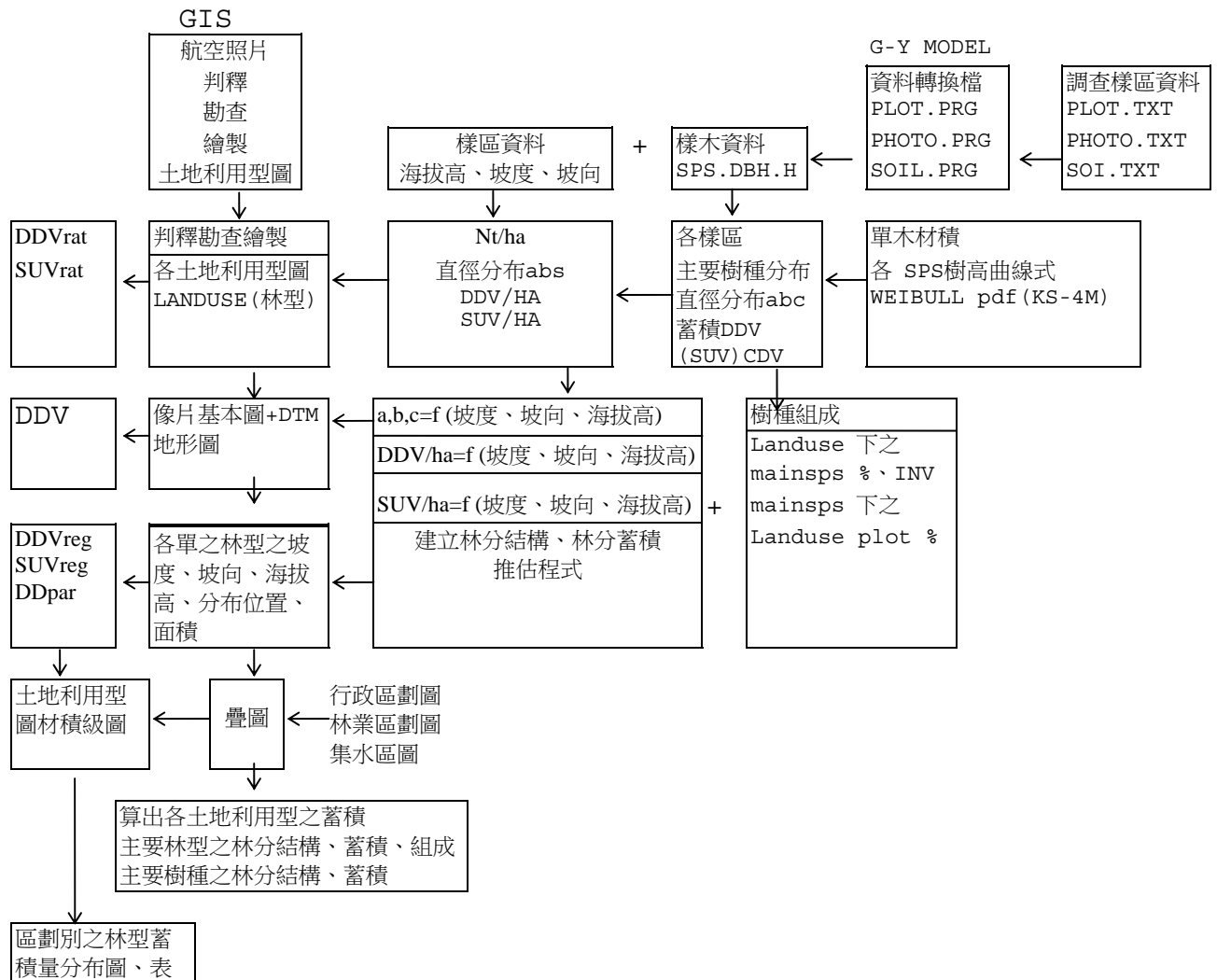
I. 以農林航空測量所所製做之數值地形模型 (DTM)，疊合以上繪置製之各種天然林林型圖求出各種天然林涵蓋的 $40m \times 40m$ 地理單元的平均坡度、坡向、海拔高。

J. 以 (E) 所建之材積迴歸推估模式，來推算 ($regression\ estimation$) 各種天然林型涵蓋地理單元之材積並加以累計。可得各

種天然林型之林木蓄積，。

- K. 以 (H) 所建之母數 (θ) 迴歸推估模式，來推算各地理單元的母數 (parameter estimation)，再藉 (d) 的直徑分布法，亦可推算天然林型涵蓋地理單元之材積並加以累計。可得各種天然林型之林木蓄積 DDVpar。
- L. 由以上 (H) (J) (K) 可以算出 SSVre，DDVre，SUVreg，DDVreg，DDVpar 五種結果。
- M. 評估選擇最合理之材積推估值。
- N. 利用地理資訊系統圖籍資料與屬性資料關聯功能，繪製材積級圖。
- O. 與林班圖、事業區圖、集水區圖等林業、行政區劃主題圖疊合，即可獲各林業、行政區劃土地利用型的各種樹種林型之林分結構及材積分布圖與數量圖。

其資料處理流程可由圖二看出：



圖二：台灣第三資森林資源調查蓄積量資料處理流程

綜合其結果可得：

(1) 林分樹種組成分析

- A. 分析調查區主要樹種、直徑級株數及於樣區中之出現率，以決定林型、樹種別。
- B. 整理各方位、海拔、經緯度下之林型、樹種別。

(2) 林分結構分析

- A. 求算各地面樣區、航照片樣區及各樹種別、林型別之胸高直徑分布與平均每公頃蓄積量。
- B. 建立各樹種別、林型別之材積推估模式。
- C. 以Weibull pdf 來配置各樣區之直徑分布，利用五種母數求解法求出Weibull 母數a、b、c。

(3) 建立分析推估材積之模式系統。

A. 單位面積蓄積量推估

(a) 合計樣區內之所有利用單木材積式計算之樣木的材積再除以樣區面積以得單位面積蓄積量之推估 (SUV/ha)。

(b) 利用直徑分布法推估單位面積之蓄積量 (DDV/ha)。

B. 建立林分材積推估模式

(a) 建立林分生長收穫隨著UTM坐標位置、海拔高、坡度、坡向等空間因子不同而改變之趨勢變化。

(b) 利用逐步迴歸建立林分材積推估模式。

C. 推算全林分材積

(a) 利用地理資訊系統 (GIS) 的疊圖分析功能將林分蓄積，結構的空間模式與數位地形 (DTM) 和土地利用型圖整合分析、推算出組成各土地利用型積結構分布圖、蓄積量分布圖。

(b) 疊合所需瞭解的行政區劃、林業區劃、集水區區劃圖層以掌握，各種區劃之樹種組成、林分結構、林分蓄積量。

(4) 繪製各種區劃之林型材積分布圖。

3、模式驗證：以第三次森林資源調查之地面樣區之林分資料檔結果驗證各主要林型之天然林林分蓄積推估模式並分析誤差來源。

七、附錄

全省尺度、林管處尺度、事業區尺度、林班尺度

1. 依土地利用型 (主要林型)、位置座標、海拔高、坡向、坡度為分層標準，將樣區分屬歸類

2. 歸屬同類樣區內之樣木，予以合併 (pooling)，計算

(1) 樹種組成：各林型主要樹種組成。

(2) 掌握林分結構 (直徑分布) — 以Wei-bull機率密度函數建立描述各樣區或各樹種之林分結構 (直徑分布) 情形，由KS適合度測驗，發現配置理想且其母數a，b，c值可解釋各直徑分布情形。更而得知各直徑級佔有株數頻度及材積分布。

(3) 蓄積量

a、關係式、材積式為單株材積林分蓄積量推估之核心。

$$H = f(D) \quad V = f(DHF)$$

b、單位面積之蓄積量可由單株材積累計單株材積、林分蓄積量 (SUV/ha)

和直徑分布法 (DDV/ha) 計算之。

c、區劃範圍之林分蓄積可分比例推估法、迴歸推估法與母數推估法三種。

3. 由GIS土地利用型圖(主要林型圖)、位置座標分區圖、海拔高圖、坡向圖、坡度圖等圖層予疊合，可算出各種分層條件下的土地利用型面積(主要林型)面積，再以樣區分析之單位面積之材積比例推算各林型之林分蓄積量。

面積及蓄積

1. 土地利用型區分

林地部份係依林地上經濟樹種之生立木中蓄積佔多數之樹種而決定。在照片上區分林型係依據各樹種樹冠所佔面之比率和樹冠大小而判斷。林分內經濟竹林佔林冠覆蓋一半以上時區分為竹林；非林地部份則依實際狀況予以區分。土地利用型用以下列三位阿拉伯數字代號表示：

代號	土地利用型	代號	土地利用型
011	冷杉天針	114	台灣杉造林
012	鐵杉天針	115	柳杉造林
013	檜木天針	116	肖楠造林
014	松類天針	119	其他針造林
015	雲杉天針	120	人針混
019	其他天針	130	人針闊混
030	天針闊混	141	相思樹造林
040	天闊純	142	楓香造林
050	天闊混	143	樟樹造林
061	桂竹林	144	光臘樹造林
062	孟宗竹林	145	台灣櫟造林
063	麻竹林	146	桐類造林
064	荊竹林	149	其他闊造林
065	綠竹林	150	人闊混
069	其他竹林	161	桂竹造林
070	天竹針混	162	孟宗竹造林
080	天竹闊混	163	麻竹造林
090	天竹針闊混	164	荊竹造林
111	檜木造林	165	綠竹造林
112	松類造林	169	其他竹林
113	杉木類造林	170	人竹針混
代號	土地利用型	代號	土地利用型
180	人竹闊混	639	其他果園
190	人竹闊針混	640	其他墾地

600	灌木林	650	伐木跡地
611	天然草生地	700	道路
612	箭竹地	710	建築用地
613	牧草地	720	苗圃用地
620	茶園	730	水田
621	甘蔗地	740	防火線
622	蔬菜地	750	工礦用地
629	其他旱作地	760	土場用地
631	香蕉園	770	墓地
632	鳳梨園	780	鹽田
633	柑橘園	790	魚塭
634	桃.李.梅園	800	其他
635	蘋果.梨.水蜜桃園	900	裸露地
636	檳榔園	930	水面(河床、溪流、水庫、池塘)

地形資料

(18) 海拔高：

從像片基本圖上讀出樣點的海拔高並填寫四位數，不足位數者左邊補“0”。例如海拔公尺則填“0100”。

(19) 坡度級：

由該照片樣點在像片基本圖上的位置之等高線，用坡度表來直接量測。其代號如下：

代 號	坡度級	坡 度	代 號	坡度級	坡 度
1	平：	0-5度	4	急：	26-35度
2	緩：	6-15度	5	險：	36-45度
3	中：	16-25度	6	絕：	>46度

(20) 坡向：

指照片樣點在像片基本圖上之大地形的坡向，共分為個方位，其代號如下：

代號	方位	代號	方位
0	平地	5	南
1	北	6	西南
2	東北	7	西
3	東	8	西北
4	東南		

地理資訊系統於分層取樣調查設計與分析上之應用
Geographic Information System Used to Stratified Sampling
Design and Analysis

1. 地理資訊系統之空間圖層資料
2. 地理資訊系統之空間分析套疊分析
3. 分層取樣調查設計
 事前分層
4. 取樣調查資料分析與展示
 事後分層

(一) 分層取樣調查設計

以往的資源調查雖名為雙重取樣，但其實只是利用簡單逢機 (sample random) 或系統取樣 (systematic field sampling) 以求得航空照片樣點與地面樣區，再分別由航空照片判釋材積和地面調查求取樣區材積以為蓄積推測之依據而已，並沒有將航測資料與地面樣區資料加以整合，以改進材積推估能力。

另外有關分層問題，假如分層 (stratification) 成功則樣區或樣點可以充分代表不同族群，若以相同的經費來加以比較，則分層取樣將可獲得較精確的材積。然分層取樣可分為事前分層與事後分層，亦即調查前、後依不同林分性質進行分層，以利調查或分析，台灣三次資源調查皆係採事後分層。

分層可由圖或像片加以取樣，然一般由航空照片來得多，整合照片樣區與地面樣區為分層取樣設計，為較省錢的工作。然據研究由蓄積級分層是無法減少推算生長收穫的變異 (variance)，即依蓄積級分層並無法改善蓄積推算值 (Maclean, 1972)；其精確性 (precision) 則完全依照取樣數量來決定。

1、分層選點

- 森林分層可依 (1) 地形 (topographical feature)
- (2) 地位級 (site classes)
 - (3) 坡度 (aspect)
 - (4) 坡向 (slope)
 - (5) 林型 (forest type)
 - (6) 密度級 (density classes)
 - (7) 材積 (volume)
 - (8) 樹高
 - (9) 年齡
 - (10) 海拔高 (altitude)
 - (11) 事業區 (working circle)
 - (12) 集水區 (watershed)

分層之基礎盡可能以推估時相同性態值為之，目前改進以地理資訊系統之不同圖層(map layers) 重疊以獲所選定之區劃範圍。

2、分層取樣

分層(stratification)簡單講即為等級的使用，應用於取樣即為在同一等級裡其個體皆相類似(級內的變異很小)，在其他等級裡則不同。若應用分層於森林資源調查上，則可以以較少的勞力、經費來調查樣本，而達到相同的正確度(accuracy)。

適當的使用分層將可以提高取樣的效率(efficient sampling)，何謂適當，可以分成三項來加以說明：

- (1) 各定義層(defining strata)一定要獨立取樣以獲各層級的平均值。各層要獨立取樣，意即各層級的分布必須於個體被選擇計算各層級平均值前即已知。在森林資源調查上有許多原則可供分層級，如地理因子(geographical factor)、林型(forest type)或林分狀況(stand condition)等。這些因子可由以往的圖籍、航空照片判釋或事先的地面調查等事先查知。目前最有效的過程為利用高品質的航空照片、高技術的像片判釋和遙感探測技術等以便獲得地理資訊系統內所需要的高品質資料來加以研判。
- (2) 各層級的平均值必須確實不同。此平均值係指樣本的平均值(arithmetic average or mean)與離均(平方)差(average squared deviation from mean or variance)、變異(variance)或均數的變異數或標準差(標準誤差)，各層級間的差異愈大，則分層取樣之效率愈較簡單隨機取樣來得高。
- (3) 各層級的觀測值分布一定要適當。決定適當與否有二法，一為比例分派(proportional allocation)，二為最適分派(optional allocation)。
 - A. 各層級內的變異(variance)皆相同，若這樣則依各層級的權重的比例予以分派，於資源調查中，各層級的面積是分派的權重依據。
 - B. 若層級的變異數不相同，比例分派仍然可用，但最適分派(optimum or Neyman allocation)則較有效率(即以較少的成本獲得想達成的正確度)。最適分派意為觀測值與(各級面積×標準偏差)乘積呈比例分佈(distributed proportional to the product of class area & standard deviation)。(註：標準偏差standard deviation = $\sqrt{\text{variance}}$)
 - C. 若觀測值其各層級的變異既不相同且不呈比例，則分層取樣對此族

群沒有效率。

由以上所提，一般地理因子(geographic commits)如鄉鎮、集水區可能無法由分層取樣獲益，林型會好些。

如何分層級，可由許多原則決定，但最有效的原則還是與欲推估事物有密切關連者。

林木資源調查(timber cruising)的目的一般為推算材積。在這種情況最有效的分層標準為減少同層級內材積變異，或增加層級間級平均值的差異。

(二) 永久樣區之取樣規劃實務

1、永久樣區之取樣規劃

以新化林場地理資訊系統資料庫所建立之圖層(馮等 1993, 陳 1993), 配合地理資訊系統套疊(overlay)、選擇(select)之功能, 將林地內具有不同地理條件之地點加以標示, 以不同之樹種、林齡、地理座標、海拔高土壤類型、坡度、坡向做為分層取樣, 規劃永久樣區之依據。規劃步驟如下所列:

(1) 資料蒐集

蒐集造林台帳、像片基本圖、航空照片、歷年研究文獻及數位地形模型(Digital Terrain Model, DTM)。

(2) 土地利用型圖(含林型圖)及林業主題圖之彙整

依陳(1993)所建之新化林場地理資訊圖檔(林場界、林班界、林道、水系、土地利用型、等高線、坡度、坡向、立體圖及土壤圖)為基礎, 轉換成相同網格大小之網格式圖檔, 土地利用型圖依不同屬性及其林型代號(tag), 查詢出單一土地利用型(含林型)。

(3) 圖形套疊及樣區選點

以林班界圖、土壤圖、坡度圖、坡向圖, 配合已建立之資料庫圖層分別與不同的樹種(人工林)或林型(天然林)圖, 加以套疊, 產生不同地理因子與土壤類型之林型圖, 再分別對應造林台帳決定林齡。至此, 即可於圖面上將相同之林型, 不同坡度、坡向和土壤類型之各種組合區域標示出來, 並於各區域中以網格板逢機決定樣區位置。

(4) 尋找樣區、推算調查費用

套疊林道圖、水系圖林班界圖即可獲知樣區之座標以及到達樣區所需路徑、地理狀況, 以便計算樣區調查時間、費用, 更可配合GPS或航照片、基本圖, 尋找出樣區正確位置。

永久樣區坡度、坡向分層之檢定：

以Bertalanffy模式預測各永久樣區之基準林齡(14 年生)胸高直徑生長量，再以進行非均衡性變方分析，結果顯示：坡度、坡向之群間皆具有顯著差異，但坡度、坡向間之交感效應即未檢定出變異，如表一。故可知永久樣區之設立上，分層取樣之考量仍有其具體意義。

表一 不同坡度、坡向樣本變方分析表

變因	DF	SS	MS	F Value
坡度	2	25.441956	12.720978	* 6.17
坡向	3	21.403898	7.134633	* 3.46
交互效應	6	20.492197	3.415366	1.66

*表差異顯著水準% (MSE=2.061746)

*DF 為分群後之自由度

以新化林場而言，樹種組成包括了大葉桃花心木、柚木、竹、龍眼、麻六甲合歡、相思樹、樟樹、黑板樹、芒果、桉樹、苦楝等，其中以大葉桃花心木最為重要，在永久樣區之設定上實應優先考量；若欲配合林場地理資訊系統之建立，則其地理因子更突顯出在取樣上的重要性。新化林場海拔高度在公尺至152公尺間，變化不大，坡度、坡向在本研究中證實對林木造成生長上的差異，可提供今後永久樣區設置時之參考。

在實際作業上，由於坡度、坡向二種地理因子無法在分層取樣中做細度的分級，本實驗坡向以方位角表示且 $^{\circ}$ 為一等級，若依Whittaker(1960)提出之方位與水分梯度之關係，如圖一；則亦可由水分梯度來探討其對林木生長之影響。

新化林場之地勢較平坦，坡度分級以度為一級。

<hr/>															
NE		ENE		E		ESE		SE		SSE		S		SSW	
NNE		N		NNW		NW		WNW		W		WSW		SW	
<hr/>															
1	2	3	4	5	6	7	8	9	10						

圖一 根據方位之水分梯度 (Wittaker 1960)

(三) 永久樣區之取樣規劃及設立

利用新化林場已建立之森林經營資料庫(陳1993)，做為大葉桃花心木永久樣區取樣規劃之依據，林型圖、土壤圖、坡度圖、坡向圖之套疊，可得不同

坡度級、坡向級、土壤類型之分層資料，提供樣區設置時之參考。舉例而言，坡度5-15度、坡向為北向，土壤類型為壤土之大葉桃花心木林型，其可供樣區設立之範圍，可經由各相關圖層之交集分析產生，再經由林班區界、林道圖之疊合，提供樣區之位置(如圖九)，並達成以坡度、坡向、土壤類型為分層依據之目的。因新化林場之坡度區分為 6 級，坡向區分為 8 級(依據第三次資源調查)，同時包含 5 種土壤類型，疊合工作極為繁瑣，應考慮各個分級間之合併或設計輔助軟體，以節省作業時間。

4. 取樣調查資料分析與展示

事後分層

(二) 方法

- 1、資料處理—將文字檔轉入III和LOTUS 以方便整理。

將圖檔建入地理資訊系統中處理、展示。

- 2、資料分析—將地面樣區調查之單株林木資料進行林分結構與林分蓄積分析、並將結果累彙送至林分資料檔。

(1) 樹種組成

各樣區由航空照片判釋及地面植生踏勘所判別繪製之林型為植群分類並進行各樣區內樣木的樹種種類分析以得樹種組成依性質將樣區歸屬為主要樹種林型。

整理主要土地利用型內所含主要樹種林型樣區之百分比天然林之森林植群之組成、林分結構、林分蓄積是受生育環境影響。

樹種組成—用相對頻度、相對密度、相對優勢度之和表示之重要值。

重要值指數(Importance Value Index, IVI)

- A. 為學派改善指數 (Density-Frequency-Dominance Index) 密度、頻度、優勢度(或覆蓋度)絕對值，指數大小沒有一定範圍，且三介量單位不一總和以代表的不合理現象。
- B. IVI為相對密度+相對頻度+相對優勢度，每個介量最大值為，植物重要值之總合恆為，若一植物社會值由一植物組成，則其重要值為300。
- C. 重要值可顯示某植物於該林型或林分中相對重要性，且兼顧密度、頻度及優勢度之綜合特性(劉、蘇1989)。

$$\text{相對頻度}\% = \frac{\text{某植物之頻度} \times 100}{\text{林分中所有植物頻度之總和}} \dots\dots\dots (1)$$

相對密度(relative density)%

$$= \frac{\text{某一植物之株數}}{\text{所有樣區內全部植物之株數}} \times 100 \dots\dots\dots (2)$$

$$\text{相對優勢度 (relative dominance) \%} \\ = \frac{\text{某一植物之優勢度 (或覆蓋度)} \times 100}{\text{所有樹種優勢度之總和 (或覆蓋度)}} \dots\dots\dots (3)$$

$$\text{重要值指數 (Importance Value Index, IVI)} \\ = \text{相對頻度} + \text{相對密度} + \text{相對優勢度} \dots\dots\dots (4)$$

(2) 林分結構

林分結構是指林地、樹種、林木各種性態值大小分布，為樹種生長習性，環境條件和經營實施的綜合結果 (Husch, and Beer 1972)。天然林的林分結構可由林木的性態值如胸高直徑、斷面積、樹高、材積大小的分布加以說明。所以直徑分布是林分結構中的一種表示方法，乃因胸高直徑資料最易獲得，故常被使用，而直徑分布以機率密度函數 (如 (5) 式) 描述得最理想 (馮1988)。

A. 各樣區內樣木的林分結構以機率密度函數 (x, θ) 來配合直徑分布，求算出的三個母數 a, b, c 之最佳推估值 $(a, b, c,)$ 。

$$f(x, \theta) = (c/b) [(x-a)/b]^{c-1} \exp[-(x-a)/b] C \dots\dots\dots (5)$$

式中的 $a \geq 0, b > 0, c > 0$

x 為胸高直徑

\exp 為指數

(3) 建立分析推估材積之模式系統推估林分蓄積

林分蓄積可由林分內單株林木材積累積，或直徑分布法 (diameter distribution method) 由各直徑級之材積累計，或全林分生長模式 (whole stand method) 由全林分材積與樹種、林齡、地位、密度、處理間的迴歸式推估而得。由於天然林林木種類複雜，林齡不易獲得，所以不加以考慮。環境因子對植生組成、結構、蓄積之影響，有層次且次序相關性存在，以地形位置 (topographic position) 及方位 (aspect) 為最上層因子，可解釋大部分植生之地域變異，稱為環境層級觀念 (Kellman 1980)。所以以坡度、坡向、海拔高等地理因子代表地位。密度則以單位面積之林木株樹表之。

A. 整理各重要樹種之單株林木材積式，樹高曲線式。

B. 利用單株林木材積式推估累計樣區內之林木材積 (SUV)。

以樣區面積除之，以獲得單位面積之林木蓄積 (SUV/ha)

$$SUV/ha = \sum V_i / ha \quad V_i = f(D, H) \dots \dots \dots (6)$$

式中的 V_i ：株林木材積
 D ：胸高直徑
 H ：樹高

C. 利用直徑分布法求算樣區之林木材積 (DDM)。

以樣區面積除之，以獲得單位面積之林木蓄積 (DDV/ha)

$$DDM/ha = Nt \int g(x) f(x, \theta) \dots \dots \dots (7)$$

式中的 Nt 為樣區內林木株數換算之單位面積之林木株數
 $g(x)$ 為胸高直徑的函數式如 $H=f(D)$ $V_i=f(D, H)$
 $f(x, \theta)$ ：的母數 θ — a, b, c

D. 將樣區內樣木所求算之 $main\ sps$ ， a ， b ， c ， Nt/ha ， SUV/ha ， DDV/ha 轉入樣區林分檔中。

E. 繪製林分結構母數 (a, b, c) 與材積母數 (SUV/ha ， DDV/ha) 與坡度 (slope)、坡向 (aspect)、海拔高 (AH)、林齡 (age) 之關係圖由此判斷林分結構、材積與環境、林齡之間的關係為直線或指數、倒數。

F. 建立各SPS土地利用型與林型、林分結構、材積與地理因子，建立樹種空間性材積推估模式。

建立林分結構、林分蓄積分析模式系統架構。

各林型 $SUV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $DDV/ha = f(\text{坡度、坡向、海拔高})$

各林型 $\theta = f(\text{坡度、坡向、海拔高})$

G. 將土地利用型歸併為種天然林林型及包含人工林、建地等其他土地利用型。依第三次資源調查所繪製之林型圖算出各種天然林涵蓋的面積以及顯示其所在位置範圍之天然林林型圖。

H. 利用 (c) (d) 所算出之單位面積之林木蓄積 (SUV/ha) 與 (DDV/ha)，配合 (G) 算出之各種天然林涵蓋的面積，可以以比率推算 (ratio estimation) 出各種天然林型之林木蓄積。

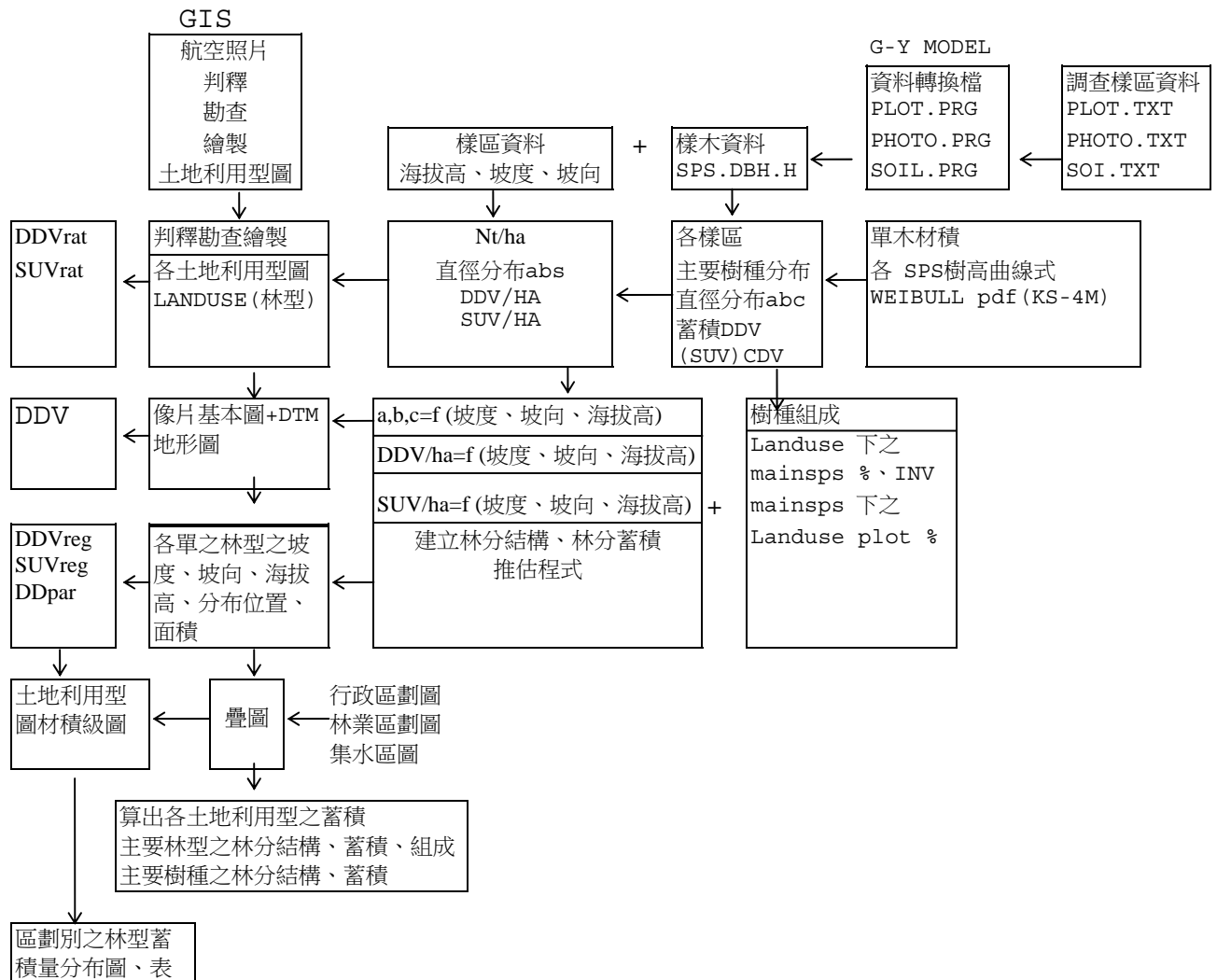
I. 以農林航空測量所所製做之數值地形模型 (DTM)，疊合以上繪置製之各種天然林林型圖求出各種天然林涵蓋的 $40m \times 40m$ 地理單元的平均坡度、坡向、海拔高。

J. 以 (E) 所建之材積迴歸推估模式，來推算 (regression

estimation) 各種天然林型涵蓋地理單元之材積並加以累計。可得各種天然林型之林木蓄積，。

- K. 以 (H) 所建之母數 (θ) 迴歸推估模式，來推算各地理單元的母數 (parameter estimation)，再藉 (d) 的直徑分布法，亦可推算天然林型涵蓋地理單元之材積並加以累計。可得各種天然林型之林木蓄積 DDVpar。
- L. 由以上 (H) (J) (K) 可以算出 SSVre，DDVre，SUVreg，DDVreg，DDVpar 五種結果。
- M. 評估選擇最合理之材積推估值。
- N. 利用地理資訊系統圖籍資料與屬性資料關聯功能，繪製材積級圖。
- O. 與林班圖、事業區圖、集水區圖等林業、行政區劃主題圖疊合，即可獲各林業、行政區劃土地利用型的各種樹種林型之林分結構及材積分布圖與數量圖。

其資料處理流程可由圖二看出：



圖二：台灣第三資森林資源調查蓄積量資料處理流程

綜合其結果可得：

(1) 林分樹種組成分析

- A. 分析調查區主要樹種、直徑級株數及於樣區中之出現率，以決定林型、樹種別。
- B. 整理各方位、海拔、經緯度下之林型、樹種別。

(2) 林分結構分析

- A. 求算各地面樣區、航照片樣區及各樹種別、林型別之胸高直徑分布與平均每公頃蓄積量。
- B. 建立各樹種別、林型別之材積推估模式。
- C. 以Weibull pdf 來配置各樣區之直徑分布，利用五種母數求解法求出Weibull 母數a、b、c。

(3) 建立分析推估材積之模式系統。

A. 單位面積蓄積量推估

(a) 合計樣區內之所有利用單木材積式計算之樣木的材積再除以樣區面積以得單位面積蓄積量之推估 (SUV/ha)。

(b) 利用直徑分布法推估單位面積之蓄積量 (DDV/ha)。

B. 建立林分材積推估模式

(a) 建立林分生長收穫隨著UTM坐標位置、海拔高、坡度、坡向等空間因子不同而改變之趨勢變化。

(b) 利用逐步迴歸建立林分材積推估模式。

C. 推算全林分材積

(a) 利用地理資訊系統 (GIS) 的疊圖分析功能將林分蓄積，結構的空間模式與數位地形 (DTM) 和土地利用型圖整合分析、推算出組成各土地利用型積結構分布圖、蓄積量分布圖。

(b) 疊合所需瞭解的行政區劃、林業區劃、集水區區劃圖層以掌握，各種區劃之樹種組成、林分結構、林分蓄積量。

(4) 繪製各種區劃之林型材積分布圖。

3、模式驗證：以第三次森林資源調查之地面樣區之林分資料檔結果驗證各主要林型之天然林林分蓄積推估模式並分析誤差來源。

七、附錄

全省尺度、林管處尺度、事業區尺度、林班尺度

1. 依土地利用型 (主要林型)、位置座標、海拔高、坡向、坡度為分層標準，將樣區分屬歸類

2. 歸屬同類樣區內之樣木，予以合併 (pooling)，計算

(1) 樹種組成：各林型主要樹種組成。

(2) 掌握林分結構 (直徑分布) — 以Wei-bull機率密度函數建立描述各樣區或各樹種之林分結構 (直徑分布) 情形，由KS適合度測驗，發現配置理想且其母數a，b，c值可解釋各直徑分布情形。更而得知各直徑級佔有株數頻度及材積分布。

(3) 蓄積量

a、關係式、材積式為單株材積林分蓄積量推估之核心。

$$H = f(D) \quad V = f(DHF)$$

b、單位面積之蓄積量可由單株材積累計單株材積、林分蓄積量 (SUV/ha)

和直徑分布法 (DDV/ha) 計算之。

c、區劃範圍之林分蓄積可分比例推估法、迴歸推估法與母數推估法三種。

3. 由GIS土地利用型圖(主要林型圖)、位置座標分區圖、海拔高圖、坡向圖、坡度圖等圖層予疊合，可算出各種分層條件下的土地利用型面積(主要林型)面積，再以樣區分析之單位面積之材積比例推算各林型之林分蓄積量。

面積及蓄積

1. 土地利用型區分

林地部份係依林地上經濟樹種之生立木中蓄積佔多數之樹種而決定。在照片上區分林型係依據各樹種樹冠所佔面之比率和樹冠大小而判斷。林分內經濟竹林佔林冠覆蓋一半以上時區分為竹林；非林地部份則依實際狀況予以區分。土地利用型用以下列三位阿拉伯數字代號表示：

代號	土地利用型	代號	土地利用型
011	冷杉天針	114	台灣杉造林
012	鐵杉天針	115	柳杉造林
013	檜木天針	116	肖楠造林
014	松類天針	119	其他針造林
015	雲杉天針	120	人針混
019	其他天針	130	人針闊混
030	天針闊混	141	相思樹造林
040	天闊純	142	楓香造林
050	天闊混	143	樟樹造林
061	桂竹林	144	光臘樹造林
062	孟宗竹林	145	台灣櫟造林
063	麻竹林	146	桐類造林
064	荊竹林	149	其他闊造林
065	綠竹林	150	人闊混
069	其他竹林	161	桂竹造林
070	天竹針混	162	孟宗竹造林
080	天竹闊混	163	麻竹造林
090	天竹針闊混	164	荊竹造林
111	檜木造林	165	綠竹造林
112	松類造林	169	其他竹林
113	杉木類造林	170	人竹針混
代號	土地利用型	代號	土地利用型
180	人竹闊混	639	其他果園
190	人竹闊針混	640	其他墾地

600	灌木林	650	伐木跡地
611	天然草生地	700	道路
612	箭竹地	710	建築用地
613	牧草地	720	苗圃用地
620	茶園	730	水田
621	甘蔗地	740	防火線
622	蔬菜地	750	工礦用地
629	其他旱作地	760	土場用地
631	香蕉園	770	墓地
632	鳳梨園	780	鹽田
633	柑橘園	790	魚塭
634	桃.李.梅園	800	其他
635	蘋果.梨.水蜜桃園	900	裸露地
636	檳榔園	930	水面(河床、溪流、水庫、池塘)

地形資料

(18) 海拔高：

從像片基本圖上讀出樣點的海拔高並填寫四位數，不足位數者左邊補“0”。例如海拔公尺則填“0100”。

(19) 坡度級：

由該照片樣點在像片基本圖上的位置之等高線，用坡度表來直接量測。其代號如下：

代 號	坡度級	坡 度	代 號	坡度級	坡 度
1	平：	0-5度	4	急：	26-35度
2	緩：	6-15度	5	險：	36-45度
3	中：	16-25度	6	絕：	>46度

(20) 坡向：

指照片樣點在像片基本圖上之大地形的坡向，共分為個方位，其代號如下：

代號	方位	代號	方位
0	平地	5	南
1	北	6	西南
2	東北	7	西
3	東	8	西北
4	東南		

國立中興大學森林學研究所八十學年度第一學期專題討論

森林資源取樣調查方法之研究
Study on Sampling Methods used in Forest Resource Inventory

研 究 生:許 榮 章
日 期:80.12.26.

[綱 要]

- 一.摘 要
- 二.前 言
- 三.取樣調查在森林資源調查資訊系統扮演角色
- 四.取樣單位之選定與調查
- 五.台灣全省三次土地利用及森林資源調查取樣方法之比較
- 六.討 論
- 七.結論與建議
- 八.參考文獻

一.摘 要:

森林的價值觀 ,經由單一之木材生產演變到多目標利用 ,其經營方法亦應隨之改變 ,以期在生態原則下 ,發揮森林資源的多目標功能 ,而在經營決策執行工作之前 ,必需有效地掌握森林資源經營的資訊 ,所以實施森林資源調查與

資料分

析成為不可或缺的一環 ,然而 ,在浩大之森林群落裡 ,面積太大 ,調查不易 ,為

節省時間及勞力 ,則必需應用取樣調查技術 ,以便在有限時間、經費下獲得適當

精確度的經營資訊 ;而取樣調查工作可分為兩部分 ,即如何選擇樣區及如何調查

樣區 ;茲綜合整理各專家學者之取樣理論與技術及全省三次土地利用及森林資源

調查取樣方法 ,提出討論 ,以作為將來整合性資源調查取樣之基礎。

二.前 言:

一般取樣調查原理 ,係居於費用最低、準確度最高之原則下進行 ,因此 ,必

需考慮因族群歧異度及等質性而異之樣區大小及形狀 ,如目標族群之種類豐富度

越大 ,則其等質性越小 ,所需取樣面積就越大 ,樣區排列型態就不同 ,以上係就

量的原理來區別取樣方法 ,同理 ,亦可以質的角度來決定取樣之不同 ,如族群種

類之優勢度指數越大 ,則其歧異度(Diversity)越小 ,取樣面積就可越小 ,取樣

方法亦隨之不同 ,若應用到林木材積調查 ,其直徑級分佈越少 ,則取樣面積應可

越小 ;本文即概括性地介紹森林資源各種取樣調查方法 ,以供森林資源調查者之

參考。

三.取樣調查在森林資源調查資訊系統扮演角色:

當經營者目標建立後 ,把所需要的資訊(包括容許誤差量及風險度)提供給

森林調查者(包括學者、專家) ,調查者針對這些所要的資訊 ,考量降低其誤差

率及未來資料的應用後 ,著手計劃調查工作 ;首先 ,應把目標族群範圍確定 ,如

某種商用林木調查 ,限制為活林木及連皮胸高直徑10公分以上者 ,否則 ,就會產

生不該調查者而調查之偏倚誤差 ,然後 ,搜集與調查目的有關的資料以確

保精密

度 ,再將族群內各元素排列成圖或表進行沙盤推演以便取樣之選擇 ,由以上獲得

之相關資訊 ,比較各種取樣方法的相對效率、複雜度、費用 ,以研判其可行性而

後決定樣區之大小、數目、位置 ,然後實施樣本調查 ,樣本調查除要降低理論性

之取樣誤差外 ,尚需減少人為之偏性誤差 ,如野外調查時以輪尺測定林木直徑之

方向、位置 ,均應詳加說明 ,並把測定到何種之有效單位予以確定 ;然後 ,把整

個工作程序編成調查手冊 ,其中包括目標摘要、族群描述、所要搜集之資料、測

定技術等 ,再依序進行訓練、講習、野外實地調查工作 ,中途若有新的經驗時 ,

應將手冊做定期修訂 ,最後 ,把調查資料處理及利用誤差最小、效果最佳之分析

方法進行資料分析 ,所得資訊以最理想方式如數學模式 ,或圖表提供經營者 ;其

流程如圖一 ,其中取樣方法的設計、使用 ,可說是森林資源調查相當重要的一環

,故特予詳加說明。

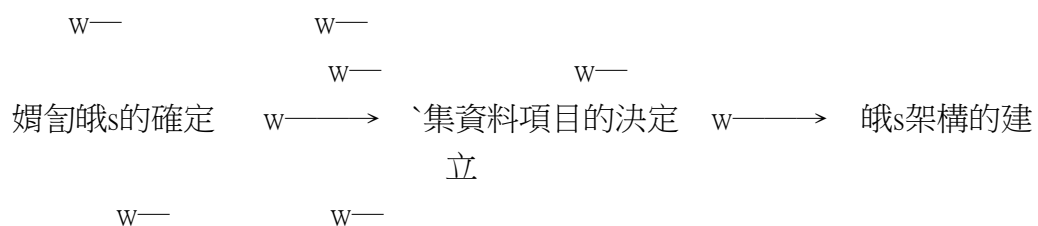
所需資訊如

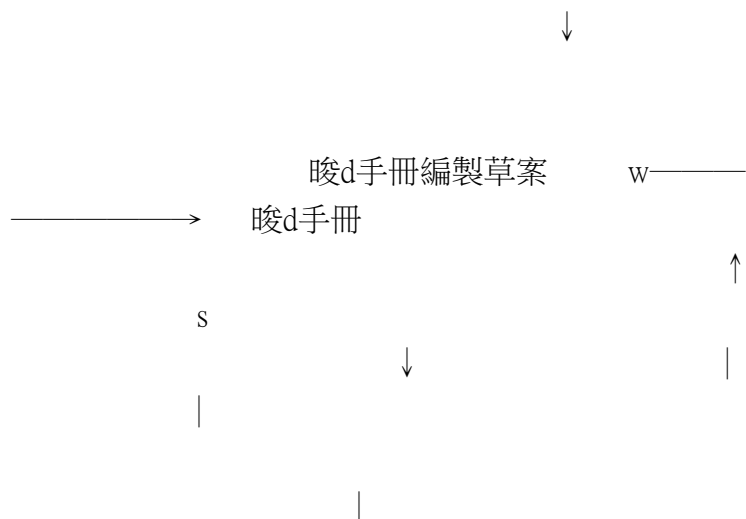
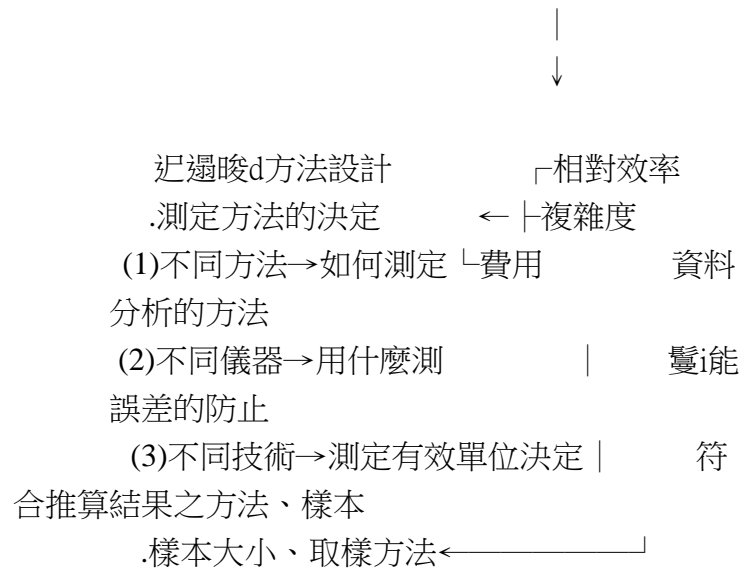
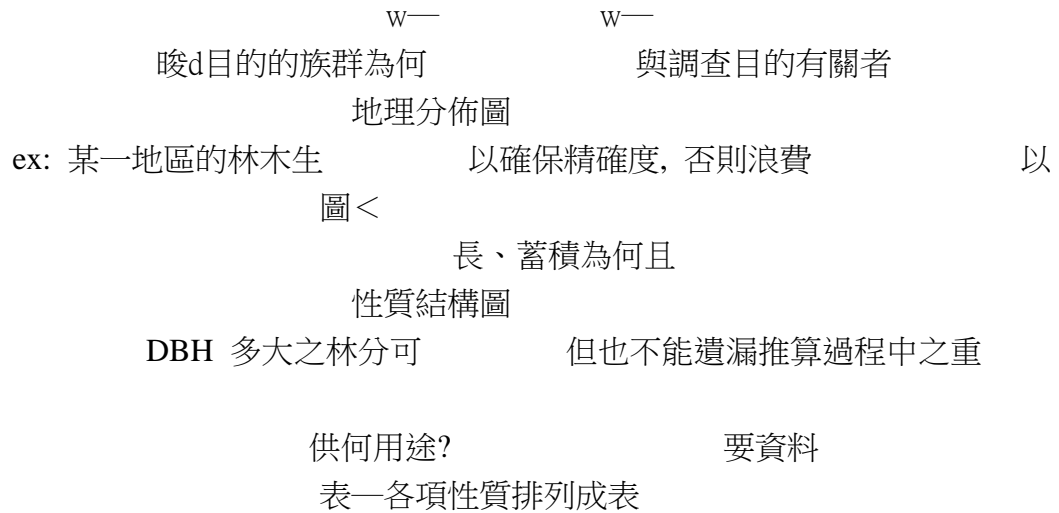
何以統計形式表之(誤差率、誤差機率)

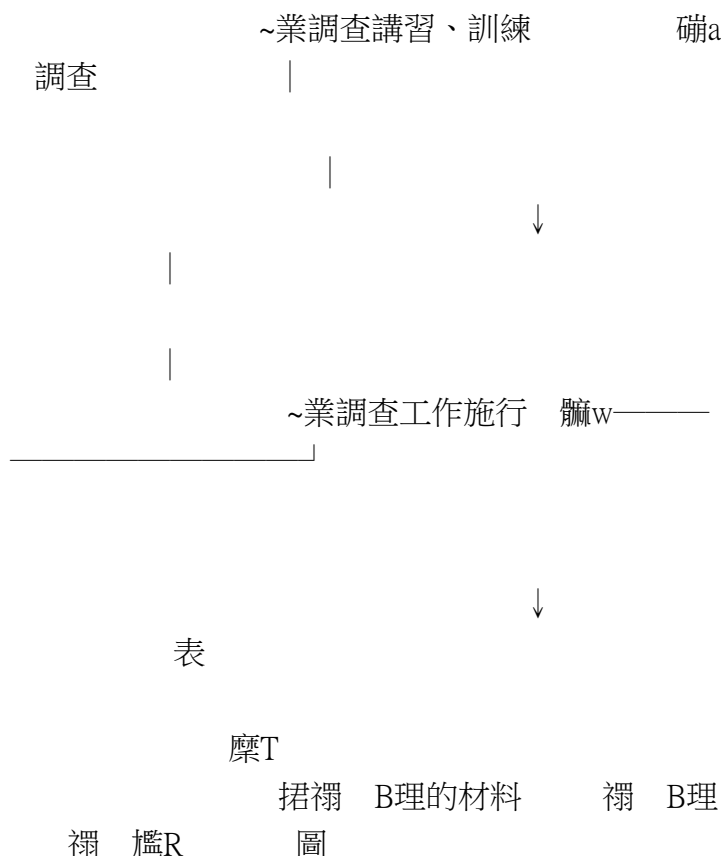
暖d目標的建立 ö

考慮未來樣

本資料的應用







圖一：調查資訊系統的設計流程

四.取樣單位之選定與調查

(一).如何決定樣區位置

要如何經營森林多項資源，首先必需獲得該項資源之有關資料，因此，取樣調查實有必要，然究竟應如何取樣，概可分為四大類：

1.逢機取樣(Random Sampling)：任何一個族群先分區編號，然後利用逢機數字表、抽籤、紙牌等方法逢機性地抽選號碼，所抽中相對區域即為取樣單位，又可分為簡單逢機取樣與分層逢機取樣(Stratified random Sampling)，是萬物取樣之基礎，符合統計分析原理，適合作變異分析與推測試驗，前者適用於形質均勻之單一林分，然若於形質不均勻之多林分時，需依樹種、材積級先予分層再逢機取樣者乃屬後者。

2.非逢機取樣(Non-random Sampling)：所選之取樣單位具有主觀意識而非任意性的，系統取樣(Sys-tematic Sampling)與選擇取樣(Selected Sampling)即為此類，而前者係逢機性任取一點為起點，每隔固定距離選擇一取樣單位，此乃薛承

健所創逢機起始之系統取樣，如全省三次森林資源調查，而後者係調查者走遍族群後認為某地可代表全體特性者，即選該地為取樣單位，在以前較粗放的林業經營，如光復前之全省森林資源調查即採用此種所謂之標準地法，或取單株調查分析之標準木法。

3.區集取樣或多段取樣(Cluster Sampling or Multi-stage Sampling)：當族群面積過大，取樣不易，為節省時間與財力，需劃分為 M 個區域，然後以逢機取樣抽取 m 個樣本組成一初級區集，再在此中依次抽取 N_i 個單位樣本組成另一個次級區集，每一個次級單位又可再分為更小單位，以此類推，可得最後樣本並予調查者謂之，惟其中以二段取樣應用最廣，如森林以林班組成初級樣本，林班中之樣區組成次級樣本，本法之最大特點在於最後取樣調查單位散佈於初級單位，並非散佈於整個族群中，可減少尋找樣區之時間，使工作集中於少數地區易於監督，在低費用下可獲得所訂精密度的推算值。

4.雙重取樣(Double Sampling or Two-phase Sampling)先於族群中以簡單逢機抽出 n 個單位組成大樣本，並測其輔助變數 X ，以求其族群均數或總和之精密推算值，然後從此大樣本 n 個單位中再以簡單逢機法抽出 m 個單位 ($m < n$) 組成副次樣本，測其主要變數 Y ，求 $Y = a + bX$ 之迴歸式以求其族群均數及總和之推算值，如應用於森林調查上則於族群相應之航空照片上，利用逢機或系統取樣法選擇照片樣點，構成照片樣本，然後在此樣本中再抽選較少之樣點組成地面樣本；此法之優點係利用較易取得之照片加以判釋，以所得資料輔助不易獲得之地面調查資料，配置較適當之迴歸式，以估計 (Estimate) 或預測 (Predict) 族群之形質，如全省三次森林資源調查，即利用此法配合分層逢機取樣求得準確之資訊。

(二).如何調查樣區

由以上種種方法可決定樣區位置，然後應如何調查樣區？可分為樣區取樣 (Plot Sampling)、和無樣區取樣 (Plotless Sampling) 兩種調查方法：

1.樣區取樣：所選取樣區具有面積、有邊際的，其形狀通常為矩形或圓形，大小則為固定或變動的，至於其個數則受族群變異數、容許誤差，及調查經費之限制而有所不同，如在 (A) 式中樣區個數係在不考慮經費下所求出的，而在 (B) 式中樣區個數則受經費、變異數之影響，但不考慮精確度，如第一次全省森林資源調查地面樣區個數即由本法求出：

(A)式

$$n = \frac{t^2(CV)^2}{(AE)^2 + \frac{t^2(CV)^2}{N}}$$

n :樣區個數
 AE :可容許誤差
 CV :變異數(可從預備調查求得)
 N :族群個數
 t :自由度 $n-1$ 之 t 表值

(B)式

$$Mo = \frac{C a}{a A + b AB}$$

Mo:地面樣區個數
C:全部經費預算
A:每一地面樣區的調查費
B:每一照片樣區的判釋費
a:面積比與材積標準差乘積之變異數
b:材積變異數

異數

b:材積變異數

樣區個數可由上述兩式求得 ,至於調查方法可從固定樣

區及變

動樣區兩方面加以介紹:

<1>.固定樣區:即樣區面積大小固定 ,一般可分為如下之樣區調查法:

A.無邊圓形樣區法 :通常用於林木材積調查(以下B、C、D 同),

即於一

林分中 ,以任一林木為中心 ,採取最近三棵木 ,計算其範圍及

胸高斷

面積 ,並推算全林材積 ,其計算式為 :

$$BA/ha = 10000 * A^3 * (Go + G1 + G2 + 1/2 G3) \text{ 或}$$

$$V = dm^2 * 1/A^3 * (a * dm^2 + b * dm + r)$$

BA:胸高斷面積和

A3:第三棵樹至圓中心林木之距離

Gi:每株斷面積(i=0...3) Go:圓中心林木斷面積

dm:林木平均直徑

a、b、r:母數

B.六木取樣法 :於林分中以任一點為中心 ,量取最近六株之直徑、距離 ,以推算林分之胸高斷面積 :

$$BA/ha = 10000 * 1/r^2 * 1/4 * (d1^2 + d2^2 + \dots + 1/2 * d6^2) \\ = 2500 * 1/r^2 * (d1^2 + d2^2 + \dots + 1/2 * d6^2)$$

r:第六棵林木之直徑

d:六棵林木至中心點之距離

C.樣木法 :組成樣區之全部個體一律調查 ,利用統計學原理 ,在

某

種水準之下可推測全林分之信賴區間。

D.帶狀橫截樣區法(Belt transect plot sampling) : 當一族群分佈於不同海拔高度 ,有帶狀分化之梯度變異時 ,樣區調查應順著

梯度、

帶狀實施 ,以期所調查之資料較能代表此種多環境變化之族群 ,本法適用於推移帶(Ecotone)之取樣調查。

E.雙截樣區法(Bisect plot sampling) :本法類似於帶狀橫截樣區法 ,惟不同之處除調查地被族群外 ,尚需調查垂直分佈之地下部分 ,如土壤之物理性質 ,根的分佈狀況

<2>.變動機率樣區取樣調查(Sampling with varying probability) :
當一族群取樣機率與其形質大小成比例時謂之 ,故該形質的大小直接影響取樣面積 ,如林木調查 ,其胸徑越大時 ,樣區面積就應越大 ,被抽選為樣木之機率亦隨之提高 ,此即所謂的 P.P.S (probability Pro-
portional to Size) ,例如目前正進行之全省第三次森林資源調查 ,若林木平均胸徑超過30公分 ,則採用 0.1公頃矩形樣區 ,30公分以下則採用0.05公頃矩形樣區 ,本法若利用水平定角儀器調查樣區時 ,最常使用水平樣點及水平樣線法 ,前者係使用林木大小決定之圓形樣區 ,而後者為使用沿著樣線 0點、10公尺點、20公尺點、30公尺點、40公尺點等五點向兩邊林木透視決定之矩形樣區 ,而其樣線應與等高線垂直 ,其設定方法即自樣區中心點(P.C)向左右兩側測得距離約 10公尺同標高之兩等高點 ,並自中心點觀測兩等高點之方位角 ,此兩方位角相加之半即為下方位角 ,以此角加上 180度即為上方位角 ,第二次森林資源調查即採用此法 ,以上兩種方法 ,取樣機率(P)均與林木直徑大小(D)成比例
即 $P = f(D^2)$ 及 $P = f(D)$,另外尚有兩種甚少使用 ,即垂直樣點 $P = f(H^2)$ 及垂直樣線法 $P = f(H)$,此兩法之取樣機率(P)均與林木高度(H)成比例;如果族群之個體所代表之價值觀相差懸殊 ,為使高價

值者被抽中為樣本之機率較大時，則稱為 3P取樣調查法

(Probability

Proportional to Prediction) ,通常用於林木標售時，讓高級林木被抽取實測的機會增多;使用本法時,需先付予每一個體一個價值級，

如、1、

2、3、4、5級，然後以歸還逢機取樣法，決定該個體是否要實測，

若規

定抽到某級之前號碼需實測時，則越高級者被抽中實測之機率就

越大(

楊汪1977);本變動機率法除P.P.S.及3P.外，還有一種稱為列表取

樣調

查法，需將族群一一按相關連數目列成表後再逢機取樣調查，本

法工作

浩繁，通常不予使用。

2.無樣區取樣調查法(Plotless Sampling) :

所選擇之樣區沒有邊際(margin)者可稱之，本法之優點在於沒

有邊

界物之糾紛，又分為下列六種：

<1>.點筐法(Point frame Method) :用點筐架上之十支測針所觸及者即

列

為調查對象，此法通常用於低矮植被之分類調查。

<2>.線截法(Line Transect) :為帶狀橫截法之變種，無寬度之帶截法，

亦

適用於低矮植被。

<3>.最近個體法(Closest individual Method) :利用樣區選擇法選取一點，然後計算該點與其最近個體之距離 r ，由下列式子，可推測每公

頃胸

高斷面積總和：

平均距離 $d = f \cdot r \cdot 1 / \text{測點總數}$ ($f > 1$ f 為常數)

每株平均面積 = d^2

密度 = $10000 \cdot 1 / d^2$ (每公頃株數)

每公頃胸高斷面積總和 = 每株平均胸高斷面積 * 密度。如下圖

<4>.最鄰近個體法(Nearest neighbor Method) :本法與上法大同小異，

惟

r 應為與該任一點最近個體之最近距離，亦即所應調查之樣本為

與該任

一點最近個體之最鄰近個體，其計算式同上。如下圖

後

近距

公頃

平均距

<5>.逢機配對法(Random pair Method) :在羅盤線上分隔數個測點，然

選定最接近測點之喬木(A)，並於相對180度之背面中找出與其最

離之另一株喬木(B)，量 A至 B之距離 r，然後可如同上法推算每

之胸高斷面積總和，本法適用於直徑大於10公分之喬木林調查。

離 $d = 0.8R * 1 / \text{測點總數}$ (R為 r之總和)，如下圖

離 r，

* 4

<6>.四象限法(Point-centred quarter Method) :本法為以羅盤線上測點
為中心分成四個象限，然後調查每一象限最接近中心之植株與距

其適用對象及計算式均同上，惟平均距離 $d = f * R * 1 / \text{測點總數}$

(f類似於 1，R為 r之總和)，如圖

五.台灣全省三次土地利用及森林資源調查取樣方法之比較

(一).第一次:民國43-45年

樣區取樣調查法 :方格板 +地形圖 +系統取樣 +雙重分層(Double Stratified) +照片樣區判釋 +地面固定圓形樣區調查。

實際步驟 :即先將全島，利用地形圖分為林地與非林地，然後將設計

好之方

上,構

定同心

均可由

格板套於其上,並將方格點針刺於地圖上,再轉刺於照片
成照片樣本,再從中系統取樣,抽出少數地面樣區,以固
圓狀之樣區,實施地面調查,以上之照片及地面樣區個數
計算式求出。

(二).第二次:民國61-66年

樣區取樣調查法:方格板+地形圖+系統取樣+分層雙重

(Stratified

Double)+照片樣區判釋+水平樣線法之矩形變動機率地面樣區調查。

實際步驟:本法不同於上法在於先將全島雙重取樣,即利用經設計好
之系統

再轉刺

圖為原

成地面

判釋及

調查,

長

尺,該

分佈方格模套在全島地形圖上,將方格點針刺在該圖上,
到照片,組成照片樣區,再以圖名為台北市五萬分一地形
點,每隔三千公尺選一樣點(此點必與照片樣點重疊)構
樣區;然後從樣區上再分為林地及非林地,分別實施照片
地面調查,此即所謂之分層雙重取樣法;本次之地面樣區
係採用水平樣線取樣法(Horizontal Line Sampling),樣區為
方形,長度固定為40公尺,樣點作中點,兩邊各延伸20公
法屬於P.P.S.之一種。

(三).第三次:民國78-82年

樣區取樣調查法:方格板+基本圖+系統取樣+雙重取樣(Double
Sampling)+照片樣區判釋+變動矩形地面樣區調查。

實際步驟:本次大致上與第二次相同,惟本次只使用雙重取樣選擇樣
區,而

之矩

地面樣區大小係隨林形、林木密度、胸徑而定,亦即P.P.S.
形樣區取樣調查法,如於天然林,每公頃30株及平均胸徑

30公分

以上者 ,採用 0.1公頃之矩形樣區 ;若於人工林 ,每公頃

30株及

平均胸徑20公分以上 ,則採用 0.05公頃矩形樣區。

六.討 論

(一).水平樣線取樣法為最新調查法 ,既經世界各國森林學者、專家極力不斷地從

理論統計上做分析檢討研究 ,並經國內實驗結果 :(廖)使用該法 ,可減少

取樣誤差及可避免邊界木糾紛等優點 ,且曾為林務局使用於第二次森林資源

調查 ,但目前進行之第三次森林資源調查為何棄之不用 ,而採用傳統式之矩

形樣區調查法 ,是否意味著該法不適合台灣使用? 若然 ,則在如此龐大的資

源調查 ,使用一種新方法之前應先做實驗評估其可行性後再使用。否則 ,調

查結果全功盡棄 ,浪費國家財源 ,而且誤導經營方針。

(二).為何第二次及第三次森林資源調查地面樣區起始點之選擇 ,均以中正紀念堂

+ (302000 2770000)為原點 ,似乎沒依照起始點選擇法 ,又該點應如何選擇

?

七.結論與建議

(一).對未來資源調查取樣上之建議 :

1.森林生態總體性觀念日益普遍 ,對於生態調查知識之需求 ,亦將日新月異 ,

希望能藉由本文所介紹森林資源取樣調查之基本方法 ,研究如何整合應用到

整個森林生態系之調查 ,如應用3P 取樣法於野生動物及牧場同時經營之整合

取樣調查(Michael 1978) ,並因應時勢潮流 ,節省時間,減少非取樣誤差及

費用 ,取樣調查應以多段取樣及雙重取樣較實用。

2.為配合GIS作生長研究 ,應多設永久樣區作連續資源調查。

3.建議國內學者專家應積極推展三重取樣(Three Phase Sampling) ,即利用

取得迅速、容易的SPOT衛星資料 ,做衛星樣點取樣作為第一次樣本 ,

從中系

統取樣組成第二次樣本即為照片樣區 ,然後再以雙重取樣法選取地面樣區 ,

應用雙重取樣迴歸式原理 ,配合地理資訊系統 (GIS)研擬以少數地面樣區實

測資料及多數之衛星樣區判釋資料作輔助變數 ,配置推測族群資料之最適迴

歸模式 ,以獲得迅速、確實的資料 ,提供給國家經建、研究教學之用。

(二).對已進行的第三次資源調查資料分析上的建議研究方向：

在以雙重取樣及系統取樣獲得的資料,應研究如何做最有效之統計分析。

八.參考文獻

- 1.林子玉 1990 森林經營研究法 國立中興大學教務處 P.112-171
- 2.陳明義 1983 植物生態學與實驗教材 國立中興大學教務處 P.17-21
- 3.陳昭明 1969 介紹幾種新的取樣方法 中華林學季刊第二卷第4期 :224-229
- 4.馮豐隆 1991 森林調查技術 國立中興大學教務處 P.15-20
- 5.楊榮啟 1980 森林測計學 黎明文化事業公司 P.191-282
- 6.楊榮啟 汪大雄 1977 3P取樣法在林木調查上之應用 台灣林業第三卷第11-12期 :7-13, 4-9
- 7.楊寶霖 1983 森林調查學 台灣商務印書館 P.241-444
- 8.劉堂瑞 蘇鴻傑 1983 森林植物生態學 台灣商務印書館 P.172-180
- 9.廖閱郎 森林資源調查上線形樣區之應用 產銷研究發展消息 P.26-35
- 10.鄭祈全 周朝富 1991 森林資源調查多層取樣設計之研究 遙感探測第14期:18-40
- 12.Behan 1990 Multiresource Forest Management Journal of Forest 88(4) p.12-24
- 11.Cochran W.G. 1977 Sampling Techniques 3ed. John Wiley & Sons p.331
- 13.Harry V. Wiant, Jr. & Edwin D. Michael. 1978 Potential of Sampling in Wildlife and Range Management in Integrated Inventories of Renewable Natural Resource : Proceedings of the Workshop Tucson Arizona 1978 General Technical Report RM-55 P.304-306
- 14.Philip G. Langley 1978 Remote Sensing in Multi-Stage ,Multi-Resource Inventories in Integrated Inventories of Renewable Natural Resource : Proceedings of the Workshop Tucson Arizona 1978 General Technical Report RM-55 P.205-208
- 15.Wayne L. Myers & Ronald L. Shelton 1980 Survey Methods for Ecosystem

management P.389-393

環境取樣學： Sampling 目錄

Sampling 這本書涵蓋基本與標準取樣設計和推估方式，加上對相當難取樣、模稜兩可、稀少性、系統性或難以偵測族群的取樣方法。Sampling 可做為科學和其他領域研究者取樣時的參考書及研究生或進一步研究課程的教科書。Sampling 可使研究者瞭解比過去傳統取樣更有效的取樣方法：對動物族群豐富度或無法估計人口的推估、預測礦物或化石燃料數量、推估稀少性疾病的流行性等。同時介紹近來具有使用潛力的取樣理論的相關發展，包括網路取樣、適應性取樣設計、不等機率取樣設計的比例和迴歸推估、空間取樣設計等。有些重要的取樣方法理論被廣泛的發展，特別在生態學、地理學、健康科學—幾乎從統計取樣理論主流中分離，如生態取樣、設計的偵測方法，地理統計學結合預測或克雷金(Kriging)方法的迴歸推估結果被記錄，以網路取樣進行流行病學的調查等。由於 Sampling 包含傳統取樣方法及近年來相關具使用潛力的取樣理論方法，使得本書較一般傳統取樣的教科書涵蓋更廣泛的取樣課題。

Sampling 分為六個部份，共二十六章。第一部分基本取樣，涵蓋簡單隨機取樣(Simple Random Sampling)、信賴區間(Confidence Intervals)、樣本大小(Sample Size)、比例推估(Ratio Estimation)及不等機率取樣(Unequal Probability Sampling)。第二部份測量資料使用最佳的方法分析，涵蓋比例推估(Ratio Estimation)、迴歸推估(Regression Estimation)、取樣的統計值(Sufficient Statistic in Sampling)、設計取樣和模式(Design and Model)。第三部份常用的取樣設計，包括分層取樣(Stratified Sampling)、系統取樣(Cluster Systematic Sampling)、多層級取樣設計(Multistage Designs)、雙重取樣設計(Double Sampling)、網路取樣(Network Sampling)。第四部份對難以捉摸的族群可偵測的方法，包括偵測和取樣(Detectability and Sampling)、樣線法和變動圓形面積取樣(Line Transects Variable Circular Plots)、捕捉和再捕捉(Capture—Recapture Sampling)、線截法(Line—Intercept Sampling)。第五部份空間取樣，包括空間預測或克雷金(Spatial Prediction or Kriging)、空間設計(Sampling Designs)、樣區形狀和觀測方法(Plot Sharps and Observational Methods)。第六部份適應取樣，涵蓋適應取樣設計(Adaptive Sampling Designs)、適應群團取樣(Adaptive Cluster Sampling)、系統和帶狀適應群團取樣(Systematic and Strip Adaptive Cluster Sampling)、分層適應群團取樣(Stratified Adaptive Cluster Sampling)。

本書特色：

- 一、前三部份敘述傳統取樣方法，使讀者熟悉、複習基本的取樣理論(第二章至第十四章)。
- 二、網路取樣(Network Sampling)，介紹以不等機率取樣方式來進行網路取樣(第十五章)。
- 三、第四部份對難以捉模的族群可偵測的方法，介紹對無法計量真值的龐大族群的偵測方式包括偵測和取樣(Detectability and Sampling)、樣線法和變動圓形面積取樣(Line Transects Variable Circular Plots)、捕捉和再捕捉(Capture—Recapture Sampling)、線截法(Line—Intercept Sampling)(第十六章至十九章)。
- 四、第五部份空間取樣，介紹空間取樣的方法設計與相關模式包括空間預測或克雷金(Spatial Prediction or Kriging)、空間設計(Sampling Designs)、樣區形狀和觀測方法(Plot Sharps and Observational Methods)(第二十章至第二十二章)。
- 五、第六部份適應取樣，引進適應取樣的主要調查方法包括適應取樣設計(Adaptive Sampling Designs)、適應群團取樣(Adaptive Cluster Sampling)、系統和帶狀適應群團取樣(Systematic and Strip Adaptive Cluster Sampling)、分層適應群團取樣(Stratified Adaptive Cluster Sampling)等介紹與傳統取樣比較(第二十三章至第二十六章)。